License Plate Detection Using Sliding Window Method

M Bagas Dewantoro (15/380918/PA/16726)
Aflah Nadhif Hammam (16/392756/PA/17060)
Arnes Respati Putri (16/392817/PA/17084)
Jehezkiel Marciel Maelissa (16/398514/PA/17475)

Department of Electronic and Computer Science
Faculty of Matemathic and Natural Science

University of Gadjah Mada
Yogyakarta, Indonesia

***Abstract*—Reading vehicles plate numbers is a technology which very useful in security and traffic control. Image processing using machine vision is the best way to make Reading vehicles plate numbers program. Edge detection is used to make gray scale image so it can more easily to read the plate numbers. Canny Edge Detection is the best method to detect the edge of the object. This Canny Edge Detection will be more effective if we are using Sliding Window method to detect the location of the object and MLP method to train the program so it can classify the plate number or not plate number without worrying with the noise and blurring.**

Keywords—Edge Detection; Multi-Layer Perceptron (MLP); Sliding Window, Canny Edge Detection.

I. INTRODUCTION

With the recent advances in intelligent transportation systems, automatic car license plate detection and recognition (LPDR) has attracted considerable research interests. It has a variety of potential applications in security and traffic control, and much work has been done on the topic of LPDR.

However, most of the existing algorithms work well either under controlled conditions or with sophisticated image capture systems. It is still a challenging task to read license plates accurately in an open environment. The difficulty lies in the extreme diversity of character patterns, such as different sizes, fonts, distortion, occlusion or blurring, and the highly complicated backgrounds, like the general text in shop boards, windows, guardrails or bricks [1].

Computer Vision can be the answer in reading license plates accurately in an open environment because Computer Vision aims to duplicate the effect of human vision by electronically perceiving and understanding an image. One of the methods in machine vision is Edge Detection [2].

Edge detection is a crucial step towards the ultimate goal of computer vision, and is an intensively researched subjects; an edge is defined by a discontinuity in gray level values. In other words, an edge is the boundary between an object and the background. The shape of edges in images depends on many parameters: The geometrical and optical properties of the object, the illumination conditions, and the noise level in the images. Edges include the most important information in the image, and can provide the information of the object's position. Edge detection is an important link in computer vision and other image processing, used in feature detection and texture analysis [2]. In plate recognition, edge detection used for detecting the numbers. In gray scale the plate numbers will be easily to be detected.

A Multi-layer perception is a feed-forward artificial neural network model that maps sets of input data onto a set of appropriate outputs. A MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. These networks have been applied to distinct areas, performing tasks such as function fitting and pattern recognition problems, by using the supervised training with an algorithm known as error back propagation. In this paper, we proposed an improved method based on MLP neural network and backpropagation algorithm for training to recognize characters and numbers in the Indian NP [3].

II. LITERATURE REVIEW

Some literature has discusses about Image processing and CNN. Some literatures explain the libraries to make edge detection more easily. The template matching method could be created by OpenCV library combine with C/C++ programing [4]. Other than that Optical Character Recognition (OCR) library can also be used for edge detection. OCR library is combine with java programming language could make a great edge detection program [5]. Both of this library are using the common Edge Detection method

Other paper explains about Edge Detection using CNN. The paper use Sparse Connectivity, Shared Weights Neural Network, Max Pooling and graphical depiction of a model which all of this is the common CNN methods [2].

III. BASIC METHOD

A. Pixel Based Edge Detection

In digital image processing, we can write an image as a set of pixels $f_{p,q}$ and an edge detection filter which detects edges with direction φ as a matrix with $w_{x,n}$, see Figure 1. We can then determine whether a pixel $f_{p,q}$ is an edge pixel or not, by looking at the pixel's neighborhood, see Figure 2, where the neighborhood has the same size as the edge detector template, say (2N+1) x (2M+1). We then calculate the discrete convolution.

$$g_{p,q} = \sum_{n=-N}^{N} \sum_{m=-M}^{M} w_{n,m} f_{p-n,q-m}$$

Where $f_{p,q}$ can be classified as an edge pixel if $g_{p,q}$ exceeds a certain threshold and is a local maximum in the direction perpendicular to φ in the image $g_{p,q}$.

$$\begin{bmatrix} w_{-N,M} & \cdot & \cdot & \cdot & w_{N,M} \\ \cdot & \cdot & & & \\ \cdot & & w_{0,0} & & \\ \cdot & & & \cdot & \\ w_{-N,-M} & \cdot & \cdot & \cdot & w_{N,-M} \end{bmatrix}$$

Fig. 1. A (2N+1) x (2M+1) template $w_{p,q}$.[2]

$$\begin{bmatrix} f_{-P,Q} & \cdot & & & \cdot & & & \cdot & f_{P,Q} \\ \cdot & \cdot & & & & & & \cdot & \\ & & f_{p-N,q+M} & \cdot & \cdot & \cdot & f_{p+N,q+M} & & \\ & & \cdot & \cdot & & & \cdot & & \\ \cdot & & \cdot & & f_{p,q} & & \cdot & & \\ & & \cdot & & & \cdot & \cdot & & \\ & & f_{p-N,q-M} & \cdot & \cdot & \cdot & f_{p+N,q-M} & & \\ \cdot & \cdot & & & & & & \cdot & \\ w_{-P,-Q} & \cdot & & & \cdot & & & \cdot & w_{P,-Q} \end{bmatrix}$$

Fig. 2. A (2P+1) x (2Q+1) image with a (2N+1) x (2M+1) neighborhood around $f_{p,q}$ [2].

B. Gaussian Filter

The Gaussian Filter is included in a linear type filter with a weight value for each pixel set in it by using the Gaussian function. The Gaussian Filter is used to blur images and remove unwanted detail and noise. The linear process in the Gaussian filter is done by multiplying each adjacent neighbor pixel and summing the result so that it gets the result for a certain coordinate point. The mechanism of the linear spatial filter is to move the center of a filter mask from 1 point to another. In each pixel, the result of the filter at that point is the sum of the multiplication of the filter coefficients and the corresponding neighbor pixels in the filter mask range.

There are 2 components to note on the Gaussian filter that is correlation and convolution. Correlation is the process of passing mask to the image. While the definition is defined as a process for obtaining pixel values based on their own pixel values, neighboring pixels and kernel matrices. In the process, the kernel will be shifted along the rows and columns of the input image used so that the new pixel value of the resulting image will be obtained. In the Gaussian filter itself, the convolution process first rotates the filter mask of 180º and then passes the image [6].

Gaussian filter has 2 types of filters: 1-dimensional Gaussian filter and 2-dimensional Gaussian filter, where the equation below as the 1-dimensional Gaussian function.

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{\left(\frac{-x^2}{2\sigma^2}\right)}$$

And the equation below as the 2-dimentional Gaussian function. The standard deviation of the distribution is expressed as σ and x and y is expressed as coordinate points (rows and columns) in image pixels. As the value of σ gets larger, then the distribution curve of the Gaussian gets wider and the peak decreases.

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{\left(\frac{-x^2-y^2}{2\sigma^2}\right)}$$

C. Canny Edge Detection

The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images. Detection of edge with low error rate, which means that the detection should accurately catch as many edges in the image as possible 1)The edge point detected from the operator should accurately localized the center of the edge. 2) A given edge in the image should only be marked once, and where possible, image noise should not create false edges [7].

D. Convolutional Neural Network

Convolutional Neural Network (CNN) are variants of MLPs, but unlike MLP which every neuron has a size one dimension, every neuron in CNN represented in two dimension [8].

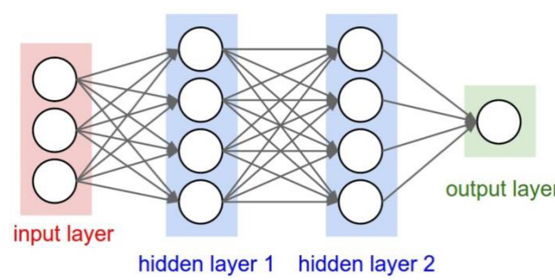


Fig. 3. Basic MLP architecture [8].

An MLP like in Figure 3. Has a layer with every layer contains neurons. MLP receive one dimension data input and propagate the data into network until produce an output. Every data input in a layer do linear operation with the value of the existing weight, and then the computation result will transform using nonlinear operation called activation function.

In CNN the data that propagate to the network is two dimensional data, so that linear operation and weight parameters in CNN are different. In CNN linear operation use convolution operation, and the weight is no longer one dimensional, but four dimension which is a collection of convolution kernels. Weight dimension in CNN are:

neuron input x *neuron output* x *height* x *width*

Because the character of convolution process, then CNN can be used only for data that has two dimension structures like image and sound [8].

CNN are inspired from biology. From Hubel and Wiesel's early work on the cat's visual cortex, we know there exists a complex arrangement of cells within the visual cortex. These cells are sensitive to small sub-regions of the input space, called a receptive field. Additionally, two basic cell types have been identified: simple cells (S) and complex cells (C). Simple cells (S) respond maximally to specific edge-like stimulus patterns within their receptive field. Complex cells (C) have larger receptive fields and are locally invariant to the exact position of the stimulus [2].

1. *Sparse Connectivity*

Fig. 4. A Sparse Connectivity template [2].

Convolutional Neural Networks exploit spatially local correlation by enforcing a local connectivity pattern between neurons of adjacent layers.

The input hidden in the m-th layer are connected to a local subset of units in the (M-1)-th layer, which have spatially contiguous receptive fields.

Imagine that layer M-1 is the retina. Units in layer m have a similar connectivity with the layer below. We say that their receptive field with respect to the layer below is also 3, but their receptive field with respect to the input is larger (it is 5

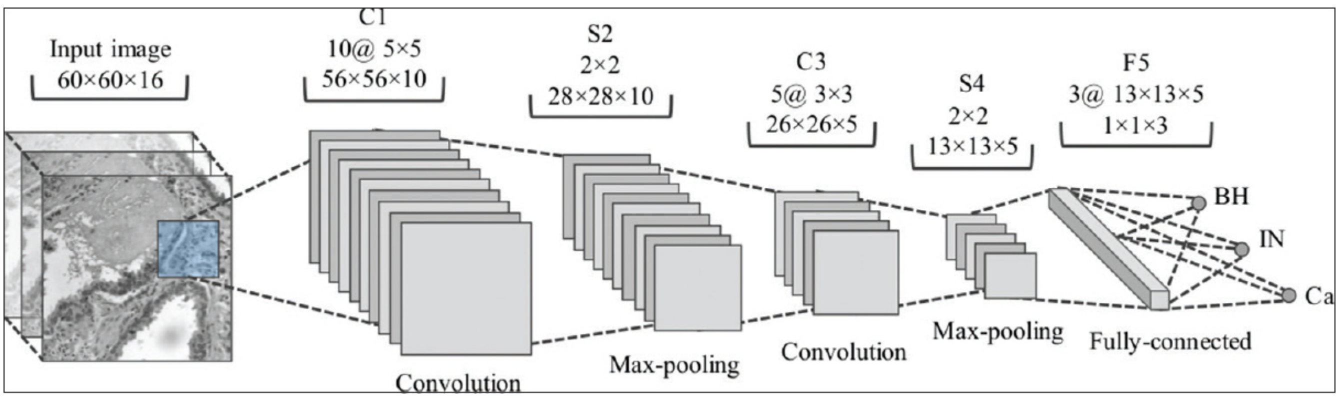


Fig. 5. Convolutional process in CNN [2].

The architecture thus confines the learnt "filters" to be a spatially local pattern, since each unit is unresponsive to variations outside of its receptive field with respect to the retina. Stacking many such layers leads to "filters" (not anymore linear) which become increasingly "global" however. For example, the unit in hidden layer M+1 can encode a nonlinear feature of width 5.

2. *Shared Weights Neural Network*

Hidden units can have shift windows to this approach results in a hidden unit that is translation invariant. Now this layer recognizes only one translation invariant feature, what can make the output layer unable to detect some desired feature. To fix this problem, we can add multiple translation invariant hidden layers:

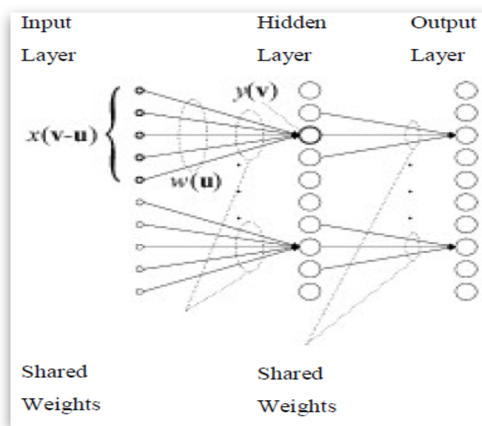


Fig. 6. Shared Weight Neural Network Edit Imager.[2]

At the learning stage, we should present the same image with shifts otherwise the edge detection would happen only in one position what was useless [2].

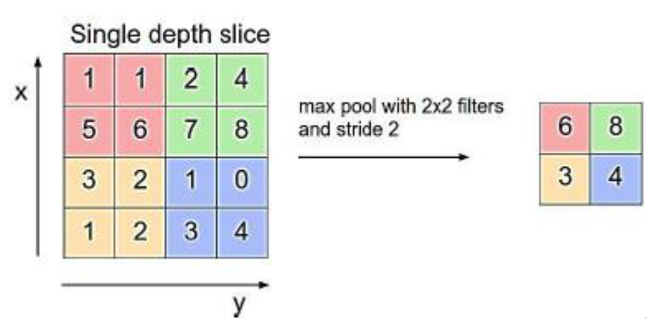
3. *Max Pooling*

Max pooling is useful in vision for two reasons:

1. *It reduce the computational complexity for upper layers.*
2. *It provides a form of translation invariance.*

If max pooling is done over a 2x2 region, 3 out of these 8 possible configurations will produce exactly the same `output at the convolutional layer. For max pooling over a 3x3 window, this jumps to 5/8.

Max pooling divide output from convolution layer into some mini grids then take the maximum value in every grid



to arrange image matrix which has been reduced. The process ensures to obtain the same features even image object has translation.

Fig. 7. Max Pooling Operation [8]

Max-Pooling layer used to reduces image size so it can easily replace with a convolution layer with a same stride as max-pooling layer [8].

IV. Implementation

The Implementation is first to collect the data set from the FMIPA UGM entrance CCTV where we get 1000 data from the camera. The next step is preprocessing where we use the Gaussian Filter and Canny edge detection before we train using sliding window to detect the object and MLP to classify the object. We use the architecture where we also apply the Relu activation layer and Sigmoid activation layer for the dense layer and for the hidden layer. Before the dense layer, we first apply the maxpooling layer with 2x2 size. The MLP use 100 epochs for the training in hope we can get high accuracy.

V. conclusion

CNN is very useful to edge detection because it makes edge detection become more accurate than any other methods in edge detection. Edge detection combine with CNN will brought image recognition technology become very effective in detecting vehicles plate number for traffic security.

For future research, edge detection using CNN method will use more specific method in CNN like YOLO method or CNN combine with Runge-Kutta method.

References

[1] H. Li, P. Wang, M. You, dan C. Shen, "Reading car license plates using deep neural networks," *Image Vis. Comput.*, vol. 72, hlm. 14–23, Apr 2018.

[2] M. A., Y. A., dan M. A., "Automated Edge Detection Using Convolutional Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 4, no. 10, 2013.

[3] S. K.M dan Sadashiva. V. Chakrasali, "AN AUTOMATED SYSTEM FOR DETECTION AND RECOGNITION OF VEHICLES NUMBER PLATE BY USING ARTIFICIAL NEURAL NETWORKS," *Int. J. Adv. Electr. POWER Syst. Inf. Technol. IJAEPSIT*, vol. 1, no. 2, hlm. 26–33, 2015.

[4] H. Mulyawan, M. Z. H. Samsono, J. Telekomunkasi, dan K. PENS-ITS, "IDENTIFIKASI DAN TRACKING OBJEK BERBASIS IMAGE PROCESSING SECARA REAL TIME," hlm. 5.

[5] D. M. Babu, K. Manvitha, M. S. Narendra, A. Swathi, dan K. P. Varma, "Vehicle Tracking Using Number Plate Recognition System," vol. 6, hlm. 4, 2015.

[6] R. D. Putra, T. W. Purboyo, dan A. L. Prasasti, "A Review of Image Enhancement Methods," vol. 12, no. 23, hlm. 8, 2017.

[7] S. Borker, "Implementation of Vehicle License Plate Recognition Using Canny Edge Detection," *Eng. Technol.*, vol. 6, no. 06, hlm. 4, 2015.

[8] I Wayan Suartika E. P, Arya Yudhi Wijaya, dan Rully Soelaiman, "Klasifikasi Citra Menggunakan Convolution." JURNAL TEKNIK ITS, 2016.