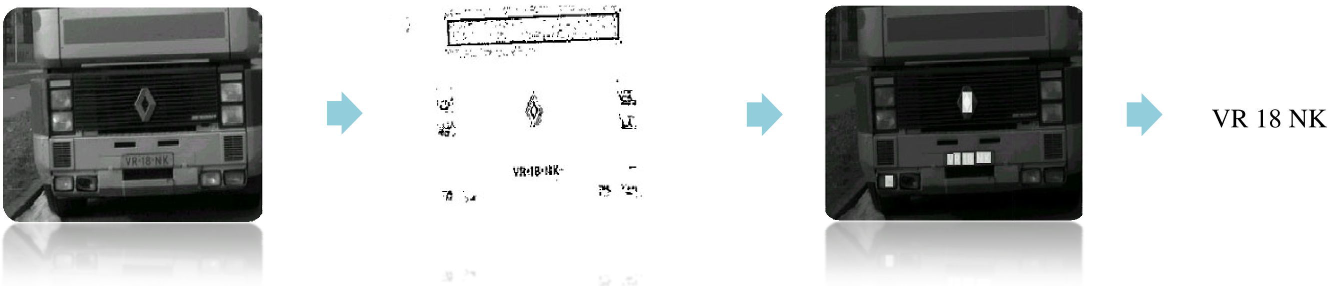
Evaluating a neocognitron type Artificial Neural Network for Car License Plate Recognition

M. Haak, S.Bos, S. Panic

Department of Man Machine Interaction
Delft University of Technology
Delft, NL



***Abstract* – In this paper, we investigate the merits and demerits of a neocognitron type artificial neural network for automated CLPR systems. In particular, we investigate a novel variation on this type, devised by Cornet. We tested the proposed CLPR system using the original software of Cornet on 14 newly acquired and three trained images. Under very strict conditions and with a semantic post-processor, the proposed CLPR system will work for very limited commercial applications. The enhanced neocognitron is very suitable for correctly recognizing isolated license plate characters. Unfortunately, correct automatic isolation of characters done by the image processor is extremely difficult (resulting in no or wrong character segmentation). The difficulty lies in the many variables such as various lighting conditions, noise or simply a bit of dirt on the license plate. The three proposed automatic image pre-processing methods work in very limited cases; images must be razor sharp and have license plate characters between 10 and 19 pixels in height. Further research is needed, but the combination with a neocognitron type ANN is certainly promising for CLPR systems.**

***Index Terms* – Neocognitron Neural Network, Car License Plate Recognition.**

I. Introduction

Car License Plate Recognition (CLPR) systems have applications in several domains. Most notable the traffic domain, where it is needed for automated tollgates. Another domain could be national security, where it could be needed for international car tracking.

In this paper we evaluate the CLPR system of engineer Bas Cornet [1], using our own test set and search for the limitations of the system and biases in the research.

We are aware of the serious ethical issues related to the application of tracking cars and that there is (much) work to do in this area. This is however outside the scope of this paper.

II. CLPR System Components

CLPR systems have three general components: 'character location/segmentation', 'character recognition' and 'high order semantics' for license plate recognition. They are briefly described below figure 1:

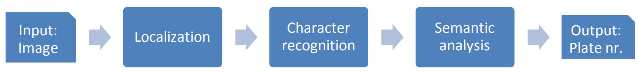


Fig. 1: Simplified flow chart of CLPR system

A. *Character location/segmentation*

Finding and segmenting characters, invariant from angle, noise (both from the digitization process as well as physical noise such as dirt), and bad illumination is notoriously hard for computers. The segmentation component described by Cornet uses two license plate features and exploit them by using filter techniques from the signal processing field, to tackle this problem. These features are:

- The color of a license plate is significantly different from the object it is attached to (car)
- The color of the license plate's background and its letters are significantly different

B. *Character recognition*

When characters are found, they need to be classified. One of the angles to approach the character recognition problem is to replicate the character recognition part of the human vision. This approach led to an Artificial Neural Network (ANN) type called 'neocognitron', invented by Prof. Fukushima[2][3] in the late 80s. The CLPR system of Cornet uses a novel variation on this architecture for the character recognition problem.

C. *High order semantics*

After all characters have been classified individually, it seems trivial to put them in the correct order to find the complete license plate number. However due to noise, individual recognition is not perfect. Problems like the classification of 'S' or '5' can be dealt with using additional information like heuristics. There are several approaches to integrate high order semantics, most incorporating domain knowledge. Examples are a syntax analyzer or a database with license plates combined with car color and type. Cornet has skipped this step in his system. In this paper we therefore only evaluate the first two components. The author does emphasize the importance of this post- processing step.

III. Set-up of Evaluation

A. *Environment*

In our pursuit to replicate the results and evaluate the proposed CLPR system, we use the original, unmodified software developed by Cornet. To convert the images conform the required format (see below) we use Adobe Photoshop version 7.0.

B. *Test set*

We acquired over 25 images from the Internet and from an Apple iPhone camera. The image content varied greatly; different license plate templates (pre and post 2008), distances, angles, illumination and with multiple cars. However, the implemented image processor requires images that are hair sharp, contain characters between 10 and 19 pixels in height and are not overexposed. Furthermore we discovered that images must not be processed by special filters. Image enhancing algorithms such as anti-aliasing or dithering, cause major problems for the implemented image analyzer. These limitations already show that this CLPR system will be inferior to human recognition. Due to the conditions discussed above, we were left with 14 images and 3 (trained) demo images.

There are also strict limitations enforced by the software so we had to manually convert all images to 8bit grey-scaled, 640x480 resolution images.

Our test set is build for two purposes; to replicate the original results and to discover some limitations of the CLPR system. The first purpose is covered by our 'normal' category, with license plate characters in the ideal height range (15-19 pixels). The second purpose is covered by the '*misc*' category. In this category we see how the system fares with images containing noise, are zoomed in or contain the post 2008 license plate.

We chose 10 percent noise as the limit a human can recognize a license plate without error. We chose 100 percent zoom, because that is the limit (19 pixels) the image processor is able to 'skeletonize' (see fig. 2) a segment without scaling.

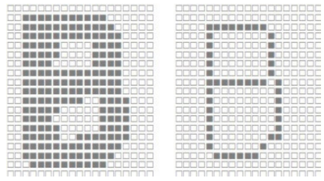


Fig. 2: skeletonize the character 'B'

We generated the noisy and zoomed images using the following procedure:

- 'Noisy' images are generated using the Gaussian distributed noise generator of Adobe Photoshop.
- 'Zoomed' images are generated by first scaling to the desired degree (using the *image size* function) and then cropped(using the *crop* tool) to 640x480.

Fig. 3: A typical image with a license plate to be recognized. a) original, b) grey-scaled, c) added 10% noise, d) zoomed in by 40%

C. *Procedure*

We started off with the '*normal*' category of the test set and continued with the '*misc*' category. The following procedure is done repeatedly (seventeen times for each of the 17 images).

- Start the application and load the selected image
- Choose automatic filter method 1 and start automatic detection.
- Store the intermediate result
- Reload the same image and do the same steps with automatic filtering methods 2 and 3
- Store the best intermediate result as the final result

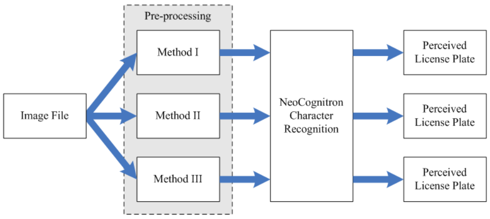


Fig. 4: Test procedure

Note: The analysis of the same image three times is necessary because there is no single method implemented that is capable of pre-processing the license plate correctly in all situations.

D. *Variables of interest*

Like in [3], we evaluate two components of the CLPR system; Character location/segmentation and Character recognition. We are interested in both full license plate recognition (if available in image) and individual character recognition. Per component we study the following variables:

Character location/segmentation:

- The amount of correctly segmented license plates (all characters found and correctly segmented, plus optional noise)
- The amount of correctly segmented license plate characters

Character recognition:

- The amount of correctly recognized license plates (all characters recognized, plus optional noise)
- The amount of correctly recognized license plate characters

E. *Measurement*

We use the same metrics as in [1] to evaluate the two CLPR system components. Although we could use different metrics, the presented metrics have two distinct advantages:

- They capture practical statistics
- They are used by the original authors, so we can compare

Character location/segmentation:

$$SR_{IP} = \sum_{i=1}^{N} \frac{\delta_i}{N}, \textbf{with } \delta_i = \begin{cases} 1 \text{ if all characters are found} \\ 0 \text{ if not all characters are found} \end{cases} \quad \textbf{(1)}$$

To calculate the success rate of the image processor SR_{IP} (PRSC in the work of Cornet), the nominator in (1) sums over all images, only counting the images where all characters of a license plate are correctly found and segmented. It is normalized by the total amount of images.

Character recognition:

$$RR_{char} = \sum_{i=1}^{N} \frac{\delta_i}{N}, \textbf{with } \delta_i = \begin{cases} 1 \text{ if character is correctly read} \\ 0 \text{ if character is not correctly read} \end{cases} \quad \textbf{(2)}$$

To calculate the recognition rate (RR), the denominator in (2) sums over all read characters, counting those correctly read. It is normalized by the total amount of correctly segmented characters for all images. It is important to realize that this is the recognition rate for the neocognitron including the skeletonizer.

$$RR_{plate} = \sum_{i=1}^{N} \frac{\delta_i}{N}, \textbf{with } \delta_i = \begin{cases} 1 \text{ if license plate is correctly read} \\ 0 \text{ if license plate is not correctly read} \end{cases} \quad \textbf{(3)}$$

The nominator in (3) sums over all read license plates, only counting the correctly read ones (with all characters plus optional noise). It is normalized by the total amount of correctly segmented license plates (with all characters plus optional noise).

IV. Results

In this section the results of the experiments with the previously mentioned variables of interest will be mentioned. The results of the original work of Cornet are presented in a column for comparison purposes.

A. *Character location/segmentation*

TABLE I
LOCALIZATION/SEGMENTATION RESULTS

	Cornet	**Our test set**
SR_{IP}	(125/144) *100% = 86.8%	(12/14) *100% = 85.7%
SR_{IP} *on char. level*	unknown	(83/85)* 100% = 97.6%

The result in table 1 seem very similar to the work of Cornet. We must add that all incorrect images lacked only one character. Furthermore, we did limited pre-processing. We only applied a sharpening filter and did one height correction. We experimented with the incorrect results and were able to enhance the images with more complex filters to achieve 100% correct segmentation. Because of the inability to use these complex filter in the original software we choose not to include those results.

B. Character recognition

TABLE II
RECOGNITION RESULTS

	Cornet	**Our test set**
RR_{char}	93.0%	(79/83) * 100% = 95.2%
RR_{plate}	72.7%	(9/12)*100 % = 75.0%

Again the result are very similar to the work of Cornet. On plate level there is a 2.3% difference. This might be caused by a numerical bias since we have just to few samples. One result stood out during this evaluation and that belonged to an image of a motor cycle license plate. It could be perfectly segmented, but only the last five characters were recognized by the neocognitron. We found out that the neocognitron was not trained to recognize the letter 'M', because its only used for motorcycle license plates and not for cars.

C. Miscellaneous

We looked into two limits of the CLPR system; its ability to segment and its ability to recognize. We also wanted to confirm our theory of perfectly recognizing the post 2008 license plates.

First the ability to segment, by looking at the SR_{IP} results when zooming in on images

TABLE III
ZOOM RESULTS

	baseline	**+40% zoom**	**+100%zoom**
SR_{IP}	100%	100%	0%
RR_{char}	100%	89%	100%
RR_{plate}	100%	33%	0%

The first thing to notice is that when increasing the license plate height by 100%, all of the trained images show large segmentation (SR_{IP}) errors. So while humans profit from the zoom, the system suffers from it, which is due to the 'skeletonizer' part of the image processor. The other figures in the table are not really interesting, they show that the characters that are segmented can still be recognized to some degree, however the full plate can never be recovered.

Next we look at the ability to recognize, by looking at the recognition rate on both character and full plate level, when adding noise.

TABLE IV
NOISE RESULTS

	baseline	**+5% noise**	**+10% noise**
SR_{IP}	100%	33%	33%
RR_{char}	100%	83%	57%
RR_{plate}	100%	0%	0%

Unfortunately, these figures show that the segmentation component also suffers from the noise, its therefore impossible to judge the results on plate level, because only one plate gets correctly segmented. On character level we see that the neocognitron fairs well when adding only a small amount of noise. This gets worse rapidly when adding more noise, however at 10% noise even humans cannot identify the characters at a glance.

Fig. 5: Image with 10% noise, license plate is PD-21-DJ

Lastly, we validated our expectation regarding recognizing the post 2008 license plates. The CLPR system was able to recognize the license plate correctly. This was expected since its font remained the same. We had to apply some extra sharpening filters for some reason to make it work, but this had probably to do with the bad image quality.

V. DISCUSSION

The '*character location/segmentation*' results are similar to the results found in the original work by Cornet. However, to make them similar we did extensive manual pre-processing work (sharpening, ensuring the license plate height lies between 10 and 19 pixels and other work) on our test set images. These labors can be omitted in the situation where an image is taken from a fixed position (so that the license plate height is correct), under ideal illumination conditions and with a good camera (taking sharp, high resolution images). The interested reader is left with the original report the read about the bottlenecks in the image processor, resulting in these heavy requirements.

The '*character recognition*' results are slightly better than the original work (which is already quite excellent), mostly because the input is highly optimized by our manual pre-processing. If however, the neocognitron is provided with slightly bad input, recognition rate deteriorates rapidly.

The strict requirement by the image processor and the lack of a post processor make the proposed CLPR system in total, less suitable for wide sense commercial applications. There is much room for improvement though. Better noise cancellation, other filter sequences and a multi resolution approach have proven to work in other fields like shape retrieval[4]. Another path to improve the image processor component is to discard the current features and find other features of the license plate and work with them.

In the '*miscellaneous*' category we see some unsurprising results concerning the limitations of the CLPR system. The neocognitron is perfectly able to detect all characters on the new license plate mask. This is expected because the new mask uses the same font as the old mask. In combination with a good post-processor the system would be able to distinguish old and new masks. Also, when zooming in on the license plate, the image processor is unable to locate the characters, despite a noiseless biniarized license plate. This is because the license plate height range is hardcoded in the image processor, possibly to limit the noise gained when skeletonizing in the next step (see original work). The last result showed us that the neocognitron is able to recognize characters with a great deal of noise (just an intrinsic feature of the neocognitron), sometimes even outperforming humans.

While investigating the report, we found few inconsistencies, which we now can prove to be just minor. We can only argue on the comparison with a '*multi layer perceptron*' (MLP) ANN, which was done under not identical conditions and certainly not with the optimal MLP architecture for recognizing license plates.

VI. Conclusion

Returning to our prime question concerning the merits and demerits of the proposed neocognitron for CLPR applications we can conclude the following:

Merits

- Excellent recognition rate (when provided correctly isolated characters)
- Able to deal with quite some noise, sometimes even outperforming humans
- Invariant to rotation and some degree of scaling
- Deterministic
- Fast

Demerits

- Heavily relies on pre-processor for character isolation, which in turn has trouble with illumination, slightly fuzzy images and dirt on the license plate
- Only works for one trained font (euro-style)
- Finding the training set is done intuitively, which is extremely laborious because of trial and error.

As a final remark, the results presented by Cornet could be replicated with an independently acquired data set. The devised variation on the neocognitron is certainly novel since the new 3 tier architecture easily outperforms the original work by Fukushima. The recognition performance of the neocognitron in comparison with other ANN types (like MLP or more advanced) is still an open question. Further research is needed to see if an ANN based CLPR system is feasible for wide sense commercial applications.

References

[1] Cornet, B. "Recognition of car license plates using neocognitron type of artificial neural network" (2001) *Master Thesis Report*, Delft University of Technology

[2] Shouno H., Fukushima K. and Okada M. "Recognition of handwritten digits in the real world by a neocognitron", (1999) *Knowledge-based intelligent techniques in character recognition*, pp. 17-47 ISBN 0-8493-9807-X

[3] Fukushima K. and Wake N. "Handwritten alphanumeric character recognition by the neocognitron", (1991) *IEEE Transcactions on neural networks VOL 2. No 3. May 1991*, pp. 355-365

[4] Liu, Z., Mitani, J., Fukui, Y., and Nishihara, S. 2008. Multiresolution wavelet analysis of shape orientation for 3d shape retrieval. In *Proceeding of the 1st ACM international Conference on Multimedia information Retrieval* (Vancouver, British Columbia, Canada, October 30 - 31, 2008). MIR '08. ACM, New York, NY, 403-410.