Modular Neural Network Classifier for Optical Occluded Character Recognition

Andries Broersma
Rijks*universiteit* Groningen
August 2001

Abstract

Optical Character Recognition (OCR) or the automatic recognition of characters is being done for quite some years now. In general, very good results have been achieved here. In practice however, various forms of distortion can cause the recognition rate to decrease drastically. One of the areas this applies to is ALPR or Automatic License Plate Recognition. For example dirt, screws or tow bars can cause the license plate and its characters to become partially obscured. Another common problem are fenders that, because of the angle in which some images are taken, cause the top of the license plate to become obscured. Because of this, the characters become occluded, which means part of their top is cut off. In most cases this results in not recognizing the characters. By building a modular neural network, which includes networks that are especially trained with occluded characters, we can increase the recognition rate drastically. It proves that by using a classifier after the first attempt to recognize these characters with the traditional network, we can improve our results on recognizing occluded characters without a loss in normal recognition.

Contents

1 Introduction to ALPR

Optical Character Recognition (OCR) has been done for many years. One of the fields this automatic recognition of characters is used in is ALPR or Automatic License Plate Recognition. ALPR is mainly used in law enforcement control and security systems. A camera is placed above or at the side of a road and for example images are taken when a vehicle crosses a red light (see figure 1-1). Normally a person would look at such images and copy the vehicle's license plate. One can imagine that this is very time consuming work. A machine could also recognise these license plates and, in this case, could even send a ticket to the owner of the vehicle involved.

Figure 1-1. ALPR is used here for law enforcement control.

In short, ALPR consists of four main steps (see figure 1-2 and 1-3). First the license plate must be located in the image and cut out. Then, out of this license plate, the individual characters must be found and isolated. The next step is the recognition of the individual characters, often referred to as Optical Character Recognition or OCR. Finally a syntax check is applied to validate whether the recognized characters form a correct license plate.

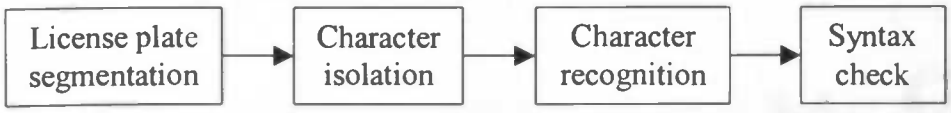


Figure 1-2. ALPR can be divided into four steps.

We will see that the problem with occluded license plates we will discuss throughout this article mainly occurs in the character recognition phase.

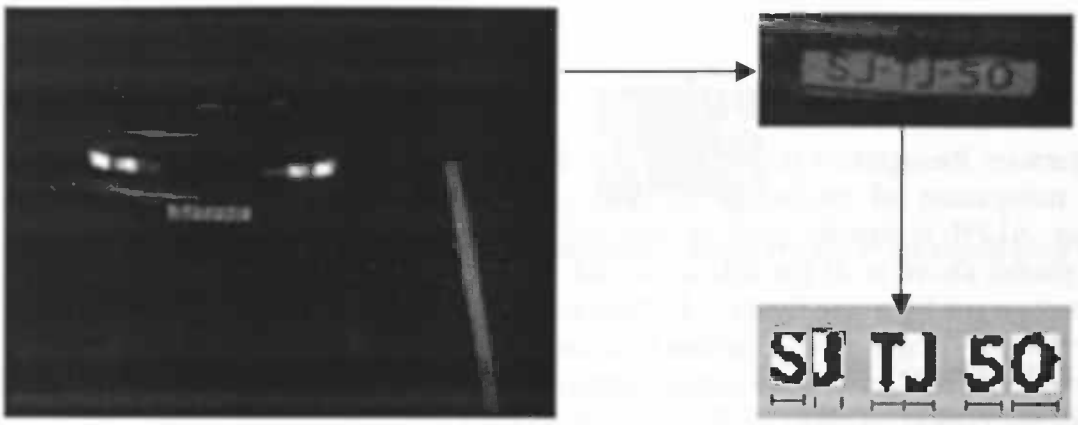


Figure 1-3. First the license plates is segmented from the image. Then the individual characters are isolated, followed by their recognition. Finally a syntax check is applied.

2 Occlusion

Although automated recognition of license plates is having very good results, the license plates that could not be recognized are often easily recognized by human. It is believed by many that no matter how good a system you build, it could only equal human in this, never surpass. Of course, some images of license plates are to dark for us to see and are no problem for the machine, but with the simple use of a filter it becomes readable for us too. Also the machine is faster and does not lose its concentration (and hence makes more mistakes) as humans do after recognizing license plates all day.

License plates that are very much inclined or upside down for that matter are not recognized by machine while for humans this causes no problem at all. Another problem, which we will be discussing throughout this article, is occlusion.

2.1 Background and foreground occlusion

When we say that a license plate is occluded we mean that the characters on the license plate cannot be seen completely. We distinguish two types of occlusion: background and foreground occlusion. Although the images are often greyscale images, the character recogniser only deals with black and white images. Everything black is treated as if belonging to the character and is called the foreground colour. The rest is white and called background colour. If background occlusion occurs then there is occlusion because of the background colour (white) obscuring some of the foreground colour. This is for example caused when an external object obscures (part) of the license plate. With foreground occlusion there is occlusion because the foreground colour (black) is obscuring some of the background colour, for example with a tow bar. So in the first case there is a part missing from the character and in the latter case there is additional foreground colour attached which now belongs to the character (see figure 2-1).

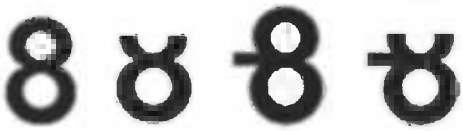


Figure 2-1. The character "8" suffering background and foreground occlusion. Both forms can also occur at the same time.

Both cases can also occur at the same time and both can cause problems. Here we focus on background occlusion.

2.2 Top occlusion

Of course in theory, background occlusion has many forms. Parts of the character can be missing at any place. From the left, right, bottom or top but also a part inside the character. In practice not all forms occur or occur often enough to be worth looking at. Sometimes occlusion of a complete license plate occurs from the left or right side but in practically all cases this causes at least the far most left or right character to become completely obscured. In that case the whole license plate can of course no longer be recognized. The occlusion we will discuss here is top occlusion (see figure 2-2). With top occlusion there is a part missing from the top of the characters. Here we only look at top occlusion where there is an equal part missing from all the characters.

Figure 2-2. Top occluded license plates that are not recognized by the current OCR.

2.3 Causes of occlusion

Occlusion of license plates, foreground as well as background, can have many causes. The causes that occur in practice can be divided into three groups.

2.3.1 Dirt

Dirt is quite common on license plates. In most cases this causes no problem because there is still enough contrast between the characters and the dirt. In some cases however, the dirt covers (part of) one or more characters, in which case they are sometimes no longer recognized. Dirt covered license plates are most common on trucks. Because dirt is mostly dark, it usually causes foreground occlusion.

2.3.2 Overlapping parts

Overlapping parts cover parts of a license plate. This can be anything from tow bars, screws, bolts, wires, chains to ladders (see figure 2-3). Screws or bolts are very common but usually cause no problems. Tow bars are less common but very often result in not recognizing the license plate. Overlapping parts in general also mostly occur on trucks. In most cases this causes foreground occlusion. Notice that in the examples given, all occluded characters are easily recognized by human.

Figure 2-3. License plates with overlapping parts.

2.3.3 Angle

In practically all cases, the top occlusion we will discuss is caused by the angle in which the photograph is taken. In many cases, for example in law enforcement control, the camera is positioned above the road. This way the photograph is always taken in an angle (see figure 2-4). Often this causes no problem. In some cases however, the fender on the back of the vehicle causes problems. If there is a fender above the license plate, or if the license plate is placed somewhat deeper in the back of the vehicle, these surrounding obstacles may causes the license plate to become occluded. Because the camera is placed above the road, this always causes top occlusion and the top is cut off in a straight line.

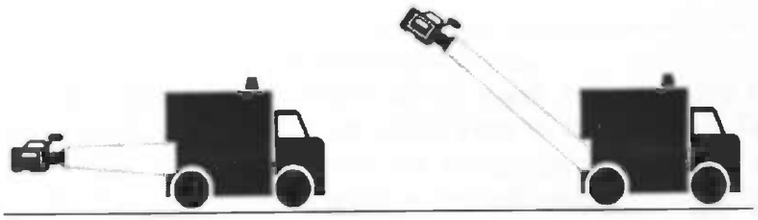


Figure 2-4. Because of the angle in which the photograph is taken, the top of the license plate may become obscured.

3 Literature survey

License plate recognition is clearly nothing new. Research in this area has been done for quite some years now. As mentioned earlier, in practice quite a few ALPR systems are being used in for example law enforcement control and surveillance systems. Also, a lot of research is done trying to improve such systems [2, 3, 7, 17, 19]. Improvement, in most cases, means improving the recognition rate. If these rates are mentioned it can be seen that they often differ a lot. The problem with these recognition rates is that they cannot easily be compared. Reason for this is that the images of the license plates that are to be recognized are often very different [13]. In some cases the images are taken at close range, directly behind or in front of the vehicle [5, 15]. In other cases the images are taken from a greater distance, which normally results in smaller license plates, which are harder to be recognized. Also, the quality of the cameras being used differs a lot. Obviously a cheap camera usually performs worse than an expensive (better) one.

3.1 Occlusion

Most articles regarding ALPR mention the fact that there are many external causes for not correctly recognizing the license plate or characters within the license plate [4, 9, 16]. In most cases these causes are dirt, screws or stickers. These objects may cause a minor form of occlusion, which usually causes little trouble. Fuzzy logic is sometimes used because of the roughness, which causes the characters to become "fuzzy" [20]. Most articles just mention these causes and do not deal with them. The fact that occlusion is quite rare is the reason for most articles not dealing with it. There are examples of systems that were tested with only a few hundred license plates of good quality. In those cases, occlusion does not even occur once.

3.2 Character isolation phase

More is said about the character isolation phase. As said, this is the phase where the license plate has been located and cut out and now the characters have to be isolated. Occlusion also causes problems here. Because of the (in this case foreground) occlusion, one or more characters can sometimes not be isolated. If we look at figure 3-1 we can see a license plate with six characters. The first two as well as the last two are not connected to each other or to the border of the license plate. They are easily isolated. The middle two in this case are connected to the border. The reason for this can be a sticker or screw for example. In many other cases such objects also cause characters to become connected to each other. Because of this the two characters cannot be isolated very easy. The solution that is used in these cases is often interpolation [10, 11, 18]. Based on the four already found characters and the characteristics of a license plate, the middle two connected characters can be located. This way, they are still isolated. Often, this only affects one or two characters and after the isolation they are still complete enough to be recognized by the characters recogniser. In this

case, occlusion is a problem in the isolation phase and not in the recognition phase. Characters that are still occluded after the segmentation are then referred to as characters that cannot be recognized.

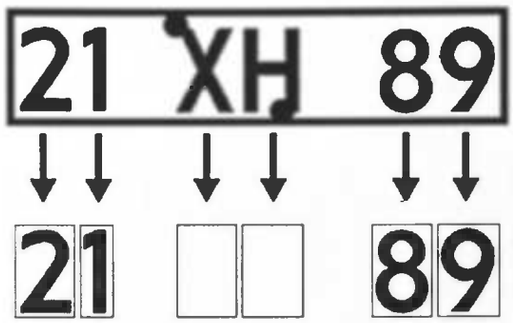


Figure 3-1. Interpolation is often used for isolating characters.

3.3 Isolation or recognition

As we read in other articles, occlusion may cause problems in the isolation phase. The type of occlusion in those cases however is one where there is not necessarily a part missing from the characters but they are simply connected, which causes difficulty separating them. In our case, the characters may also be connected, but they are also missing a part from the top. This way, isolating them may be a problem because they are connected, but now also problems arise when the occluded isolated characters are passed through the character recogniser. When we say in general that the occluded license plates are not recognized, this is more due to not recognizing the individual isolated characters then to not being able to isolate them. We used a test set of license plates that had some sort of occlusion and isolated the characters from it. Figure 3-2 shows the results. Six or more segments were found in 69 percent of the license plates. Notice that Dutch license plates contain six characters. If more segments are found they are usually one or more lines on the left or right side of the license plate. In practically all cases they are easily discarded. In the other 31 percent of the plates, not all six characters were properly isolated. These cases are thus a problem of the isolation phase. Notice that if for example only one segment was isolated, this not necessarily means that one character is isolated. In most cases all six characters were then isolated as one segment. When two, three four or five segments are isolated, it is often the case that one or more characters are isolated as one segment. The test set we used contained license plates with all sorts of occlusion. It was noticed that with the top occluded license plates, the percentage of correctly isolating all six characters was even higher. Because we focus on such license plates here, we will treat top occlusion as a problem that is mainly caused by the character recognition phase.

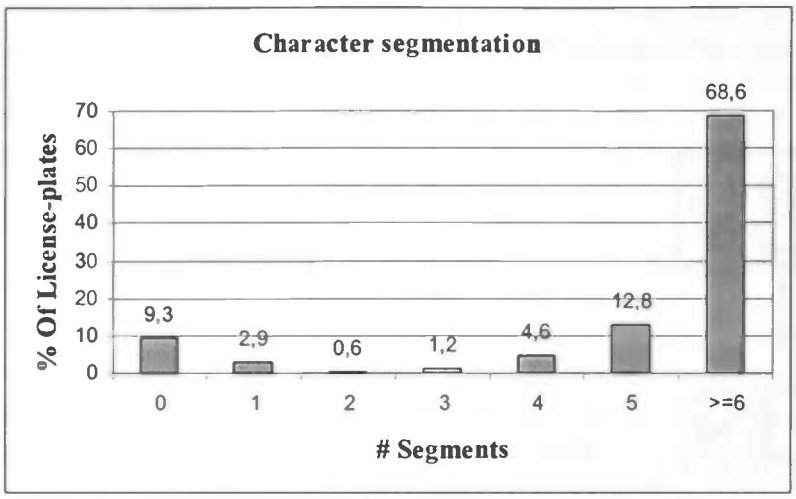


Figure 3-2. The number of segments that were isolated from the license plates.

Figure 3-3 and 3-4 show the graphical output of the character isolation. In figure 3-3, characters were isolated from an occluded license plate. Despite the occlusion, all six characters were properly isolated. Figure 3-4 shows that sometimes not all characters can be isolated. In this case, we see our example when the middle two characters are isolated as one segment.

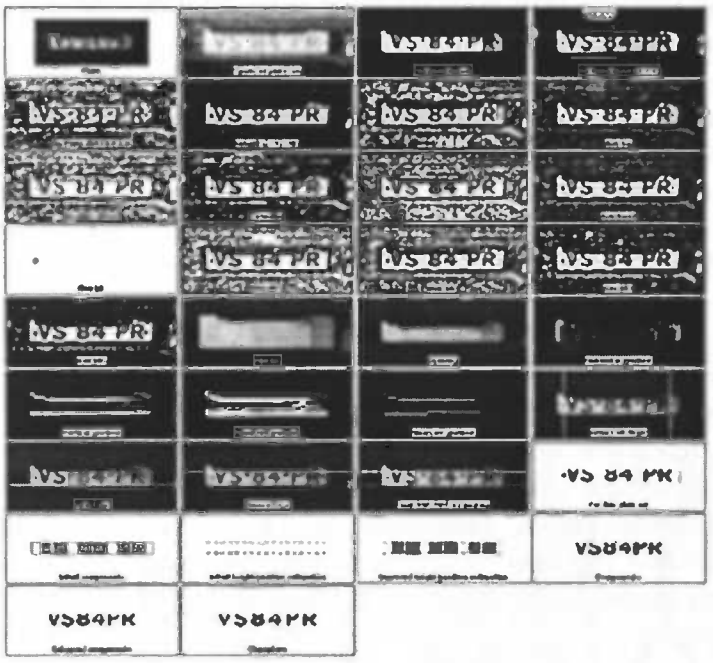


Figure 3-3. Despite the occlusion, all six characters are properly isolated.

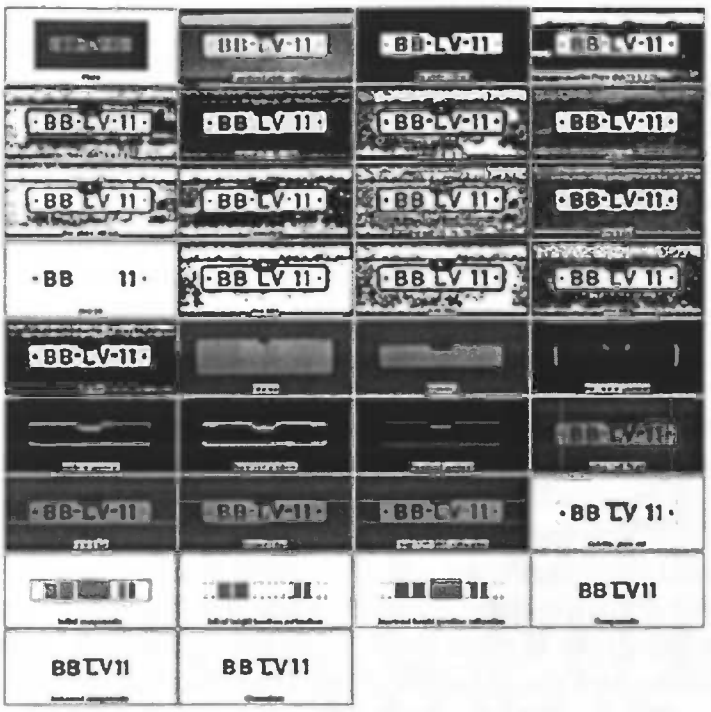


Figure 3-4. In some cases, not all characters are properly isolated. In this case the middle two.

4 Percentages occlusion

Before really dealing with occlusion, it is important to know the amount of occlusion that occurs in practice. In theory, the amount of occlusion can range from 0% to 100%, with 0% being a license plate with no occlusion and 100% being a totally occluded licence plate. To get a good feeling about the percentages in practice we looked at 175 real life images of license plates that were occluded and therefore not recognized by our ALPR-system now. Figure 4-1 shows some examples of this set. It can be seen that the amount of occlusion ranges a bit. Compare the top row with about 15 percent occlusion and the bottom row with some 30 percent occlusion. Although in theory occlusion starts at 0% percent, in practice it seemed that it started at about 10 to 15 percent. There are two reasons for this. First, as we will see later on, license plates with occlusion below 15 percent are not a problem or at least a much smaller one for the character recogniser. The second reason is that, as said earlier, the reason for this type of occlusion is, usually, the fender of the vehicle. In practice, this fender causes between 15 and 30 percent occlusion. This way, as was verified by looking at the percentages of the set occluded license plates, it is concluded that the occlusion in practice ranges from 15 to 30 percent. The average and most common amount of occlusion was about 20-25 percent.

Figure 4-1. License plates with different amounts of occlusion.

4.1 15 percent border

Figure 4-2 and 4-4 show the effect of 15 and 30 percent occlusion on the character templates. These templates are the characters that fit the font used on Dutch license plates. The double images of the same characters indicate the old and new font of that particular character. The dot in each image indicates the centre mass. If we look at the case of 30 percent occlusion, we can see that most characters are still easily recognized by human. The exceptions are the **1**, **I** and **T**. There is no way they can be distinguished anymore. If there is no knowledge of the fact that these characters are occluded, the three characters are surely all identified by human as being an **I**. Characters such as the **J** or **L** on the other hand have intuitively hardly changed. All other occluded characters are fairly easily distinguished and thus recognized correctly by human. We conclude that some occluded characters cause more problems than others for humans to recognize. We will see later on that this is also true for the character recogniser. If we look at each character as a shape with its own specific characteristics, we see that occluding the character can alter these characteristics. The alternation of some characteristics however has more effect than others, both intuitively as in practice. If we look at the **H**, its vertical lines above and below its horizontal centre line are equal of length. If we occlude this character for 30 percent, the upper vertical lines become shorter than its lower vertical lines and its horizontal line is no longer centred. If we occlude the **7** for 30 percent it completely loses its horizontal line at the top. In both cases the characteristics of the characters change, but intuitively the first case is not as big a problem as the latter. We see that many characters lose a "critical" top part if they are occluded. Also many characters with a closed top part, for example the **0**, **8** or **P**, become characters with an open top part if they are occluded. This effect starts at about 15 percent. If we look at figure 4-2 again, we see that the top part of characters such as the **5**, **E**, **F** or **T** is just obscured and characters such as the **9**, **D** or **P** become open at the top. Intuitively this seems a critical amount of occlusion and we will see later on that the recognition rate of the character recogniser decreases rapidly from that point.

Figure 4-3 and 4-5 show licence plates with exactly 15 and 30 percent occlusion and the segments that were isolated from the plates. Notice that recognizing the initial plate causes no problems at all for human.

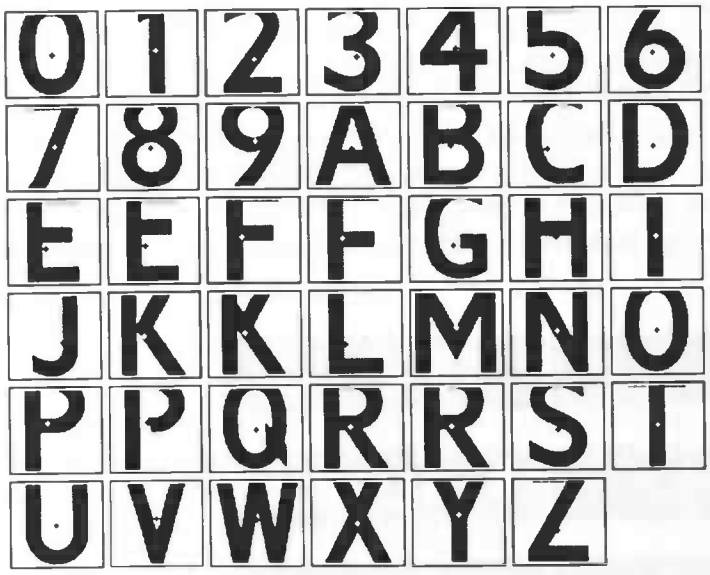


Figure 4-2. Character templates with 15% occlusion.

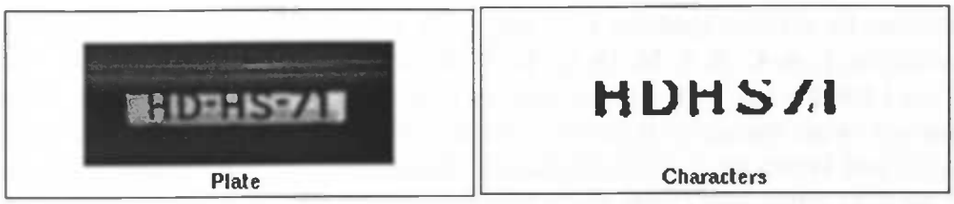


Figure 4-3. License plate with 15% occlusion and its six isolated segments.

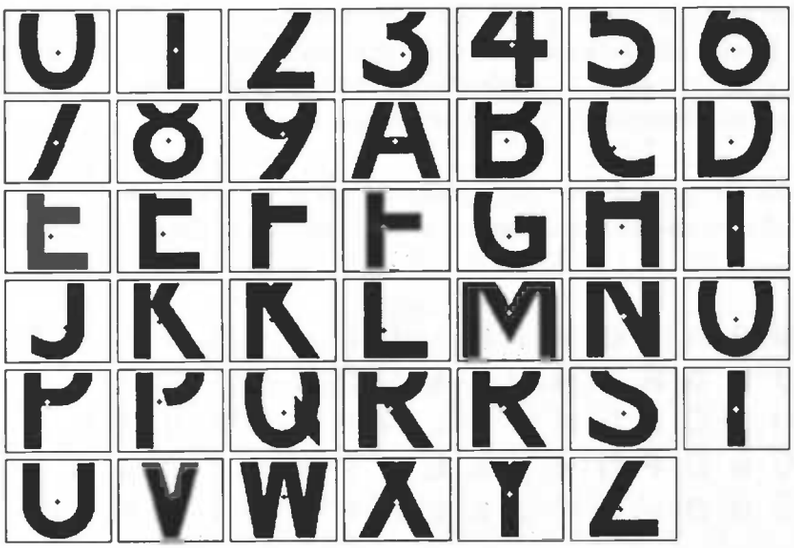


Figure 4-4. Character templates with 30% occlusion.

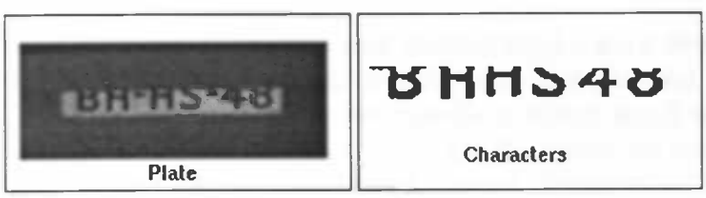


Figure 4-5. License plate with 30% occlusion and its six isolated segments.

5 Character recognition

The input of the total ALPR-system is (often) a (black and white) image of a vehicle and its surroundings (see figure 1-1 or 1-3). In this occlusion research we look at just the character recognition phase. The recognition rates that are mentioned here are therefore much higher than the recognition rates of the complete system. Character recognition is done here by using neural networks. Normally also template matching is being used, but for the simplicity of this research it is left out. If a character is passed through the character recogniser, it can be correctly recognized, rejected or falsely recognized. Of course, worst case is falsely recognizing.

5.1 Character set

The character set that is used here consists of *TIF* images. Each character is centred in a single black and white *TIF* image, which is 100x100 pixels large (see figure 5-1). The initial character set used here was acquired from the RRT1, RRT2 and RRT3 sets and consisted of some 140,000 characters, already isolated from over 23,000 Dutch license plates. Although

there are specifications for all 36 characters (0-9 and A-Z), only 26 characters are used here. For multiple reasons, the **1, A, C, E, I, M, O, Q, U, W** are left out. One of the reasons is that some characters, for example the **U**, are never used in practice. Out of this initial set, about 5,000 unfit characters were manually removed. These unfit characters consisted of badly segmented, occluded and falsely recognized characters. Because we are specifically interested in occlusion and want to artificially create occluded characters later on, we must make sure the initial characters are not occluded already. Out of these remaining 135,000 characters, a total of 88,000 (equally divided) were randomly selected. This set was then divided in two sets of 44,000 characters. One used for training and the other for testing. The total test set again was then divided into 5 smaller test sets of 8,800 characters each. Reason for this is that by using more (smaller) test sets instead of a single (large) one, we can compare the test results and look at their differences. Notice that no cross validation is used.

We will use these five test sets throughout our entire research. Later on we want to use test sets, which contain occluded characters. To create such test sets, we will artificially cut a part of the top of the character.

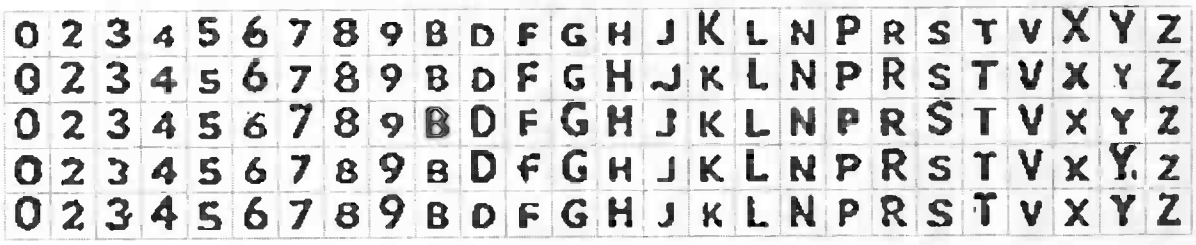


Figure 5-1. *TIF* images of the 26 different characters.

5.2 Scaling

Of course, when characters are isolated from a license plate, they can be of many dimensions depending on the pixel size of the license plate. Before these character images are passed through the character recogniser, they are up scaled or downscaled to fixed dimension.

5.3 Neural variables

Al the networks used here were trained with the following settings:

Learning epochs	50
Learning rate	0.8 - 0.2 – 0.01
Hidden neurons	40
Momentum	0.2

Figure 5-2. Neural network settings.

Normally more than 50 learning epochs are used. Because training networks is very time consuming and the goal of this project is not to get the best networks but to look at the differences of the networks and the relative gain that can be achieved, the number of learning epochs for all networks was set to 50. If we look at the learning curves (see figure 5-3) for the four networks, it can be seen that the percentage correctly trained characters stabilizes enough for our purposes around 50 epochs.

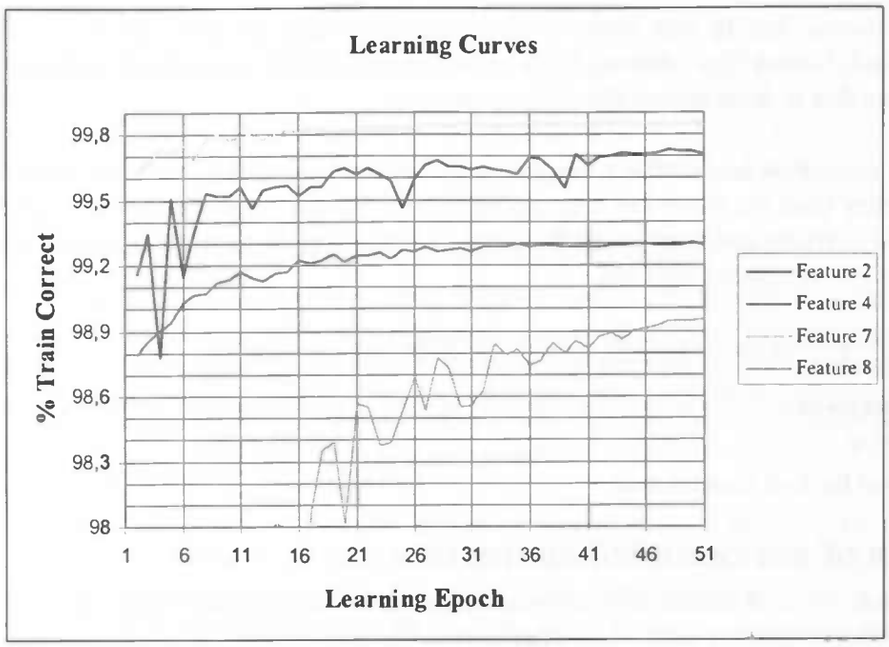


Figure 5-3. Learning curves training four no-occlusion networks.

5.4 Classes

Some characters are very similar in shape. In some cases, this similarity is so great that it was decided to combine these into a single class. The **(0, D)**, **(2, Z)**, **(5, S)** and **(8, B)** are therefore no longer distinguished from each other. Notice however that the pairs consist of a number and letter. Whether the number of letter is recognized is based purely on the syntax.

Having created four pairs, the 26 characters are divided amongst 22 classes. The train set therefore has 2,000 characters per class. The five test sets have 400 characters per class. Notice that because of the fact that for example the **0** and **D** are combined and the **3** is not, there are twice as much **3**s as there are **D**s. There are two exceptions. The **K** and **Y** were less frequent than the other characters. Instead of decreasing the total set to keep an equal amount of characters per class, it was decided to leave these two classes smaller. The train set of the **K** class now consists of 1,454 and the test sets of 290 characters. The train set of **Y** consists of 1,063 and the test sets of 212 characters.

5.5 Neural networks and features

As said, for the moment we do not look at the recognition rates of the complete ALPR-system, but specifically at the rates achieved by only the character recogniser. The character recogniser however also consists of multiple components with their own recognition rates. Based on these components, the OCR gives its final answer. One of the components is the template matcher. As said earlier, it is left out in this research. The four components used here are four neural networks referred to as feature 2, 4, 7 and 8 from now on. Figure 5-4 shows an overview of the different features with a brief explanation. In the beginning, we will mainly work with the recognition rates of these features. As mentioned earlier, characters passed through the complete character recognized can be correctly recognized, rejected or falsely recognized. Based on the neural network outputs of the four features, the OCR decides whether to accept or reject the character. How this is done is described in more detail later on. If we talk about the recognition rates of the four features, we look at only their output neurons. The character is recognized as being the character (or class) that has the highest

output neuron. This means that in this stage, a character can either be correctly of falsely recognized, not rejected. Notice that later on, both correctly and falsely recognized characters could be rejected. How this is done is described also later on.

We look at the effect occlusion has on the recognition by the individual features. We will see that one feature is better than the other but they all react to occlusion in the same way. Later we will combine these features and work with the results of the complete character recogniser (with the exception of the template matcher).

Feature 2	Connected components and horizontal / vertical pixel projection	30 inputs
Feature 4	Side view	34 inputs
Feature 7	Zoning to 7x6 pixels	42 inputs
Feature 8	Contour tracing	64 inputs

Figure 5-4. An overview of the four features used.

5.6 Recognition of not-occluded characters

To get a good reference, we first looked at the recognition rate of not-occluded characters. As said, our initial test sets consist of only such characters. We expect this to get a very high recognition rate. Figure 5-5 shows the results. The confusion matrix is read as follows: vertically are the 22 classes that were projected to the horizontally read characters and the "?", which indicates the character was rejected. The grey boxes are the correctly projected classes. The numbers in the other fields indicate the number of false projections to that particular class. It can be seen that very few errors are made and although some classes are more often falsely projected than others there are no extremes. For these not-occluded characters, the recognition rate is 99.7 percent, an error of 0.2 percent and a rejection rate of 0.1 percent. Notice that these rates are of the complete OCR. Only the results of one out of five test sets are shown. All test sets were used in most experiments but because it proved there was an extreme similarity between them, they are omitted in some stages.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400																							1
2:Z	400							3																
3	400																							1
4	400																							
5:S	400																							
6	400																							
7	400		2							1														
8:B	400	1																						1
9	400								1															
F	400																							
G	400																							
H	400														1									
J	400			1																				2
K	290	1			1											1								1
L	400																							
N	400																							
P	400																							
R	400												1		1		1							
T	400		1								1													

V	400																1							
X	400																							
Y	212																			1				
	8502	2	3	1	1	0	0	3	1	1	1	0	1	0	2	1	2	0	0	1	0	0	0	6

Figure 5-5. With no occlusion, the recognition rate of the OCR is very high.

6 Recognition of occluded characters

Now that we have a good reference, we can look at effect occlusion has on the recognition of characters. First we will use character templates. These are artificially created "perfect" images of characters. Then we will use our test sets, which contain "real life" characters. We shall compare both results and look at their differences and resemblances.

6.1 Character templates

If we look at figure 6-1 we can see the effect of several amounts of occlusion on different character templates. If we look at the characteristics of a character, it can be seen that with for example 15 percent occlusion, the **9** loses its "characteristic" closed circle at the top. This circle is now open. With the **H**, the top lines become somewhat shorter, but it loses no specific characteristic. In the case of the **9**, at 30 percent occlusion, it starts to look like a **Y**. If we look at the **Z**, at 30 percent occlusion its shape is not similar with any not-occluded character. Even for human, it would be a serious guess between a **2** and **Z**. Intuitively the occluded **R** also could cause great problems. As soon as its closed circle becomes open, its characteristics become very similar with a **K**. If we test this theory in practice we can see that indeed for example the **9** is rejected at 15 percent and recognized as a **Y** at 25 percent occlusion. On the other hand, the **H** is not rejected or falsely recognized until 30 percent occlusion. We can thus say that not only the amount of occlusion has effect on the recognition rate, but also the character that is involved. The results of figure 6-1 are based on the complete character recogniser including the template matcher.

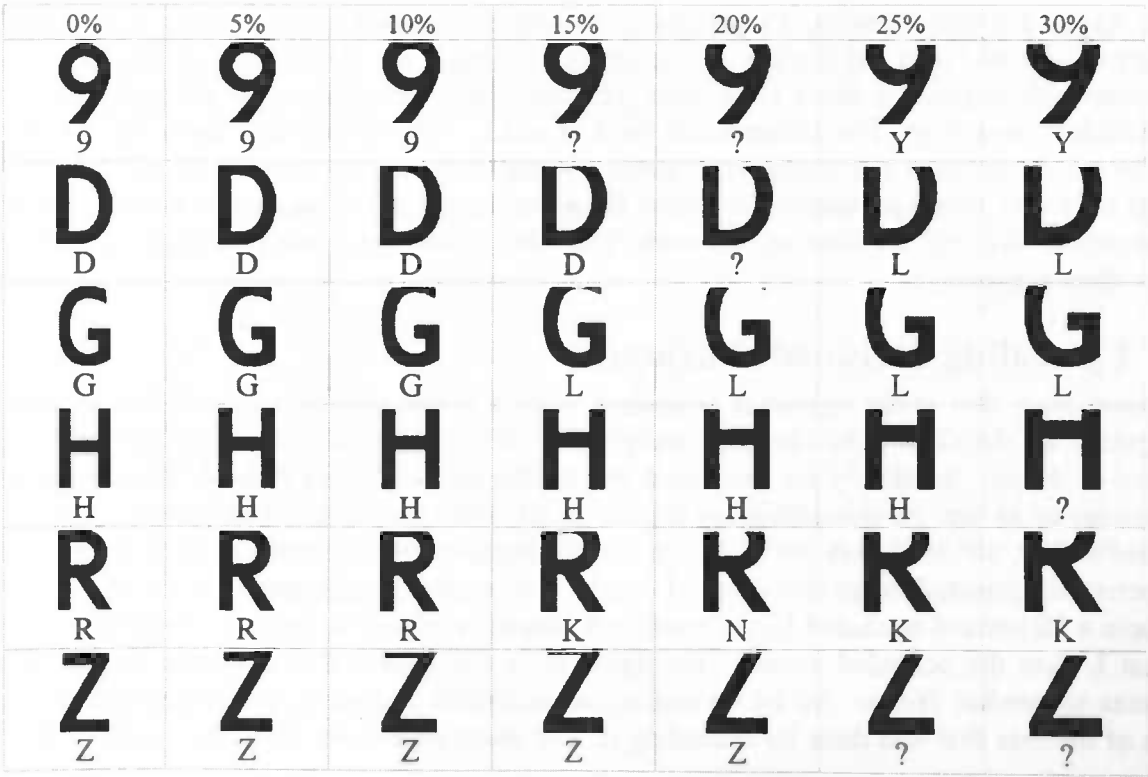
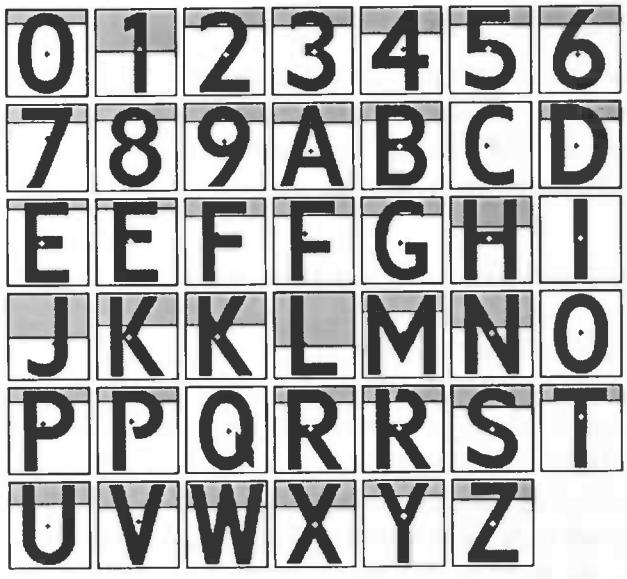


Figure 6-1. Some characters are more sensitive to occlusion than others.

Now that we have seen that both character type and amount of occlusion have an effect on the recognition, we will look at all characters individually and determine the point of occlusion where the recogniser rejects or falsely recognizes the character. This is done with the complete OCR, including the template matcher. Figure 6-2 shows the results.



0	15%
1	51%
2	16%
3	15%
4	29%
5	17%
6	19%
7	15%
8	13%
9	15%
A	18%
B	15%
D	16%
E	8%
F	15%
G	15%

H	31%
J	58%
K	37%
L	63%
M	19%
N	37%
P	10%
R	15%
S	20%
T	15%
U	19%
V	28%
W	26%
X	25%
Y	13%
Z	22%

Figure 6-2. The red line indicates the maximum amount of occlusion that the complete character recogniser can handle before rejecting of falsely recognizing the character. Notice that again, about 15 percent is a critical point.

Some characters such as the **1** and **A** are also mentioned here. They are ignored throughout the rest of this article but since they are recognized by the complete OCR, they were not left out here. The **C**, **I**, **O** and **Q** were not recognized by the OCR and are therefore left out of the table. As we explained earlier, 15 percent occlusion is a critical point. At this point, many characters become "open" at the top. As we see in figure 6-2, about two third of the characters have their critical point at about 15 percent. The other characters have their critical point at a much higher percentage. The extremes are the **1**, **J** and **L**. We will see the reason for this later on. The results in figure 6-2 confirm our statement that occlusion up to about 10 or 15 percent causes only very minor problems. Of course these results are based on perfect templates with no distortion. We will see later on that with "real life" characters, more problems occur even below these percentages.

6.2 Up scaling occluded characters

We have seen that some character templates with a large amount of occlusion are still recognized by the OCR. The character template **L** for example is still recognized up to 63 percent occlusion. Intuitively we explained this by the fact that even if the **L** misses a large percentage of its top, its characteristics remain intact. Only the vertical line becomes shorter. As said earlier, all characters are scaled to fixed dimension before being passed through the character recogniser. This is also done of course with occluded characters. If we upscale for example a 50 percent occluded **L**, the result is a character with a far greater resemblance to a normal **L** than the occluded version (see figure 6-3). Only now, the horizontal bottom line becomes somewhat thicker. So by up scaling an occluded character, we sometimes restore much of the loss that was done by occluding it. The same applies for characters such as the **J** or **1**.

Normal L 50% occluded L Up scaled L

Figure 6-3. Because the occluded character is up scaled, its resemblances with the original character increases.

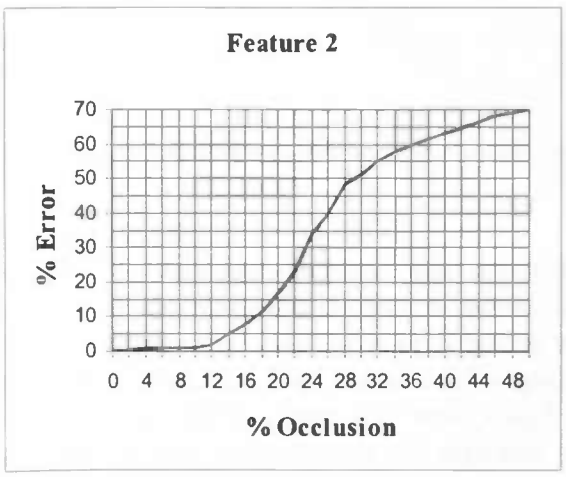
6.3 Real life characters

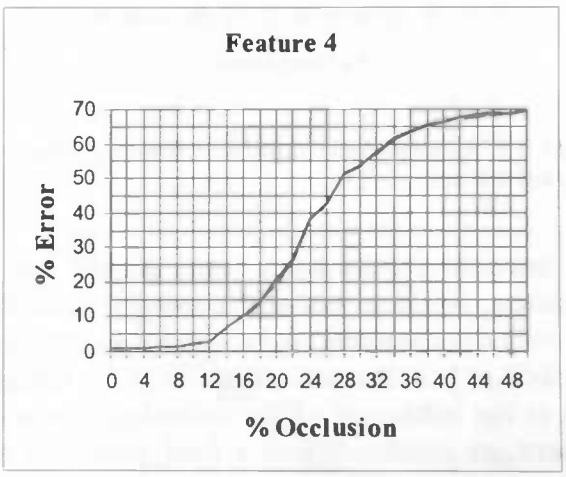
Now that we have seen the effect occlusion has on the characters templates, we will use our real life test sets (see figure 5-1) and look at the effects here. First we look at how the error increases as we increase the occlusion. We will do this first by looking at the individual four features.

It is expected for the error to increases as the occlusion increases. Figure 6-3 shows the results. As expected, to about 10 to 15 percent occlusion, the error increases relatively slow. This complies to our results with the templates. After about 15 percent the error increases more rapidly. With the templates, we saw that occlusion up to 15 percent causes practically no problems. Here we see that even below this percentage the error slowly increases. There are a few reasons for this. First, because we are dealing with "real life" characters here, there can (and quite often is) a distortion which effects the effect occlusion can have. Also, the reason 15 percent is the critical point was because with the templates, at 15 percent the characters lost their top horizontal line or became open at the top. With real life characters, the thickness of this top part is not fixed. If this line is thinner, it becomes open or lost more quickly which results in not recognising the character even sooner.

On the other hand, while most of the character templates were no longer correctly recognized by the complete OCR above 15 percent occlusion, many of the real life occluded characters are still correctly recognized by the individual features.

Although throughout this article we deal with occlusion up to 30 percent, we show results here up to 50 percent. The five different coloured lines indicate the five different test sets used. Notice the extreme similarity between them.





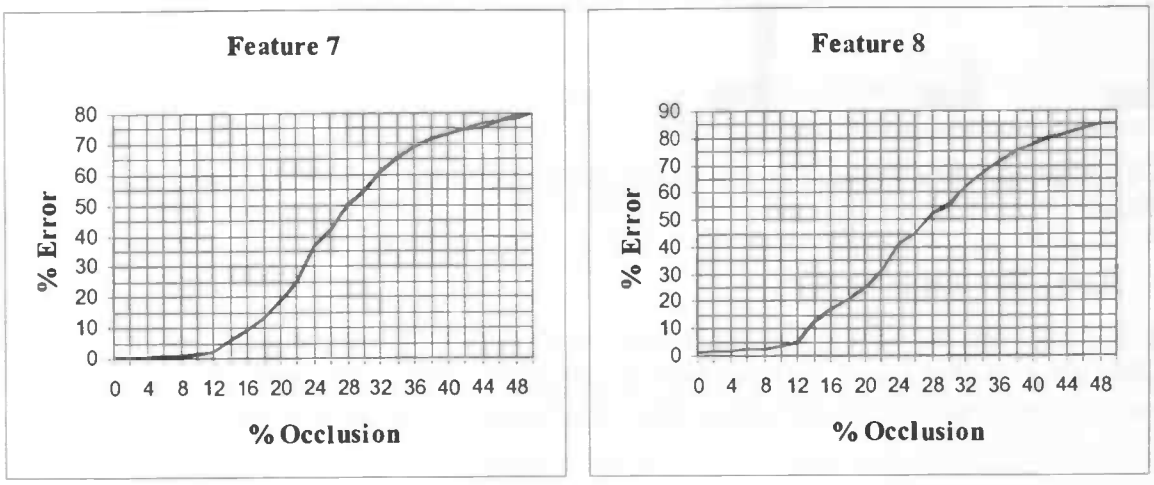


Figure 6-4. The error increases if we increase the occlusion. The five different coloured lines indicate the five test sets used.

If we compare the four different features, we see that the effect of occlusion is quite similar. All features handle occlusion equally well (or worse). However, while not-occluded characters are best recognized by feature 7, occluded characters are best recognized by feature 2. If we look at the error at 22 percent occlusion, which we stated as the average occlusion in practice, we see that in about 20 percent of the cases, the correct class did not have the highest output neuron.

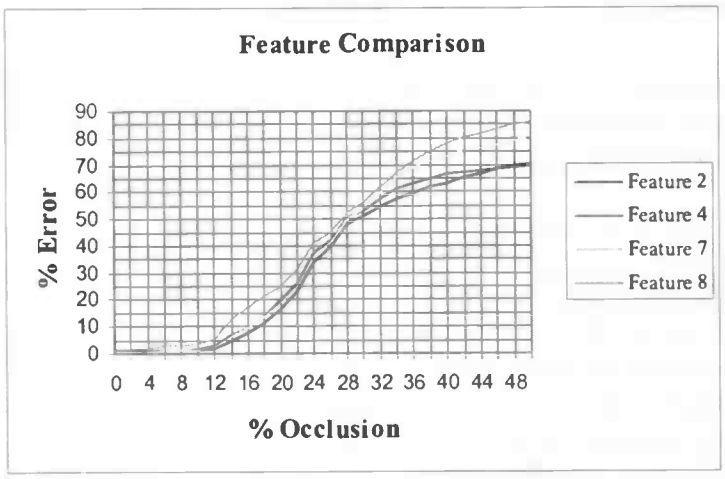


Figure 6-5. A comparison of the four features. The graphs are averages of the previous five graphs of the different test sets.

The previous results were based on the four features and were averages of all different characters. Nothing was said about the behaviour of the individual character classes. As we saw with the templates, some characters are more sensitive to occlusion than others. Now we will look at how the occlusion affects the recognition of characters by the complete OCR and look at the behaviour of the individual classes. Again, how the results of the four different features are combined into a final answer is not relevant for the moment and is explained later.

Figure 6-6 and figure 6-7 show the confusion matrices for two amounts of occlusion: the average percentage of 22 percent and our maximum of 30 percent.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400												1								6			1
2:Z	400						2	3								2						2		3
3	400				2		1							78										11
4	400									38													1	1
5:S	400	2		14																				4
6	400	153							2															4
7	400		2								1			1							1		1	4
8:B	400	55											3								2			4
9	400							6													30		123	11
F	400							15							1			37					36	25
G	400	10					8				1					4					1			30
H	400														1		2							1
J	400	1		2						1														
K	290											1				1	1		2	1			1	2
L	400																							
N	400												1											1
P	400							11			1												6	2
R	400												8		321		29							6
T	400		1					2						1		18							1	101
V	400																1						36	
X	400							2							5				3					24
Y	212																			205				
	8502	221	3	16	2	0	11	39	2	39	3	1	13	80	328	25	33	37	5	206	40	2	205	235

Figure 6-6. At 22% occlusion, many characters are falsely recognized or rejected.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400												2				1				33			7
2:Z	400				4		12	3								8						3		7
3	400	1	1		8		2	1						285							1			12
4	400									219							1			1			7	10
5:S	400	3	4	112				1																58
6	400	383																	1					4
7	400		1								4			4						3	2	1	4	6
8:B	400	262								1		2	1								1			20
9	400							21										1			28		331	6
F	400							57							1			212					29	29
G	400	127					11						5		1	16	1				9			110
H	400	3															18		1					33
J	400	1		2				1																
K	290										2	2				2	1		11	2				7
L	400																							1
N	400									1			5						1					7
P	400							107			3												5	10
R	400												2		379		11							4
T	400				1			3						2		24							2	186
V	400																1			4			226	
X	400		2		15		4	1							3				21	1				137
Y	212													1						208	1			1
	8502	780	8	114	28	0	29	195	0	221	9	4	15	292	384	50	34	213	35	219	75	4	604	655

Figure 6-7. At 30% occlusion, almost 50% of the characters are not recognized correctly.

We saw that at above 15 percent occlusion, most character templates were no longer correctly recognized, while most of the real life characters were still recognized with such occlusion. Now that we have passed our occluded characters through our complete OCR without the template matcher, many characters are still correctly recognized above 15 percent occlusion. The template matcher must therefore really decrease the performance of the OCR if we are dealing with occlusion.

If we look at both figure 6-6 and figure 6-7, we see that the results match our intuition. With 22 percent occlusion the **6**, **9**, **R**, **T** and **Y** are often falsely projected to the **0**:**D**, **Y**, **K**, **?** and **T** class respectively. Notice that the occluded **T** is projected to the **?** because we ignored the **1**. At 30 percent occlusion, the error increases even more. Now also large false projections occur that are intuitively less obvious. The false projection of the **P** to the **7** is an example of this. This occurs in 25 percent of the cases. Also the projection of the **V** to the **Y**. Intuitively one might say an occluded **V** has little resemblance with a normal **Y**. Again, because of the scaling there is a resemblance indeed. Because the occluded **V** stretches, its lower part becomes more of a standing line (see figure 6-8).

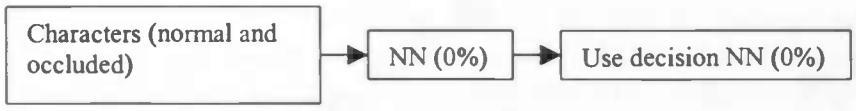
Normal V 30% occluded V Up scaled V Normal Y

Figure 6-8. Up scaling an occluded "V" increases its resemblance with a normal "Y".

The recognition rates of the OCR are 81.8% correct, 2.8% reject and 15.4% error at 22 percent occlusion. At 30 percent occlusion the recognition is 53.3% correct, 7.7% reject and 39.0% error. In comparison, with no occlusion the error was only 0.1%.

7 Training with occlusion

All the research in occluded characters and their recognition leads to a simple conclusion. The classical neural network cannot recognize such characters with a reasonable recognition rate. The reason for this is that our four different neural networks were all trained with not-occluded characters. This seems of course very logical because of the fact that occluded license plates are quite uncommon. So what happens now is that the characters are all passed through a neural network that was trained with not-occluded characters and the decisions are based on that network only. We treat the four combined features as one neural network here. In other words, not-occluded characters are being "compared" with not-occluded characters, but for example 20 percent occluded characters are also "compared" with not-occluded characters.

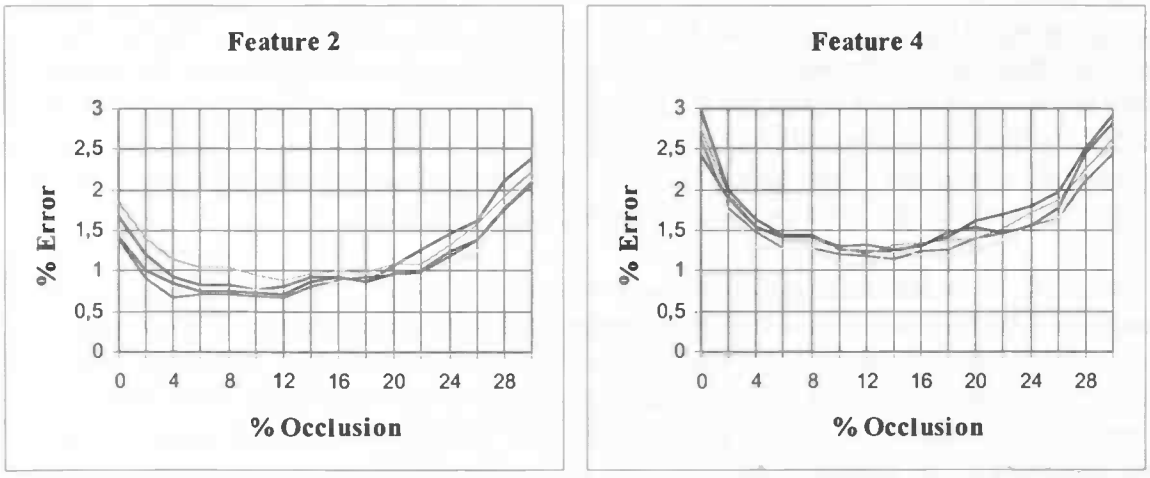


7.1 Mixed neural network

The first logical step is not to train our network (again read four features) with only normal characters, but also include occluded characters. This can be done of course in many ways.

Knowing that occlusion is very rare, one could decide to let occluded characters be only a small percentage of the total train set. If for example occlusion occurs in 1 percent of the cases, the percentage occluded characters should also be 1 percent of the total train set. This however has no effect. The characteristics of the few occluded characters are completely overruled by the far greater group of not-occluded characters and there is no improvement in the recognition of occluded characters. The solution we propose here is one where all percentages of occlusion are equally divided amongst the train set. This means that the neural network was trained with an equal amount of not-occluded characters as well as for example 10 percent occlusion. Again, we set the maximum occlusion percentage at 30 percent. The total number of characters in the train set remains the same.

By creating this neural network, we will see a tremendous improvement in the recognition of occluded characters. Our goal is of course to do so, but not at the expense of the recognition of not-occluded characters. This is however what happens if we use this method. Although the recognition of occluded characters will improve greatly, the recognition of not-occluded characters will slightly worsen. Although this worsening is very small, the improvement with occluded characters cannot compensate this loss. Figure 7-1 shows the results of the four different neural networks.



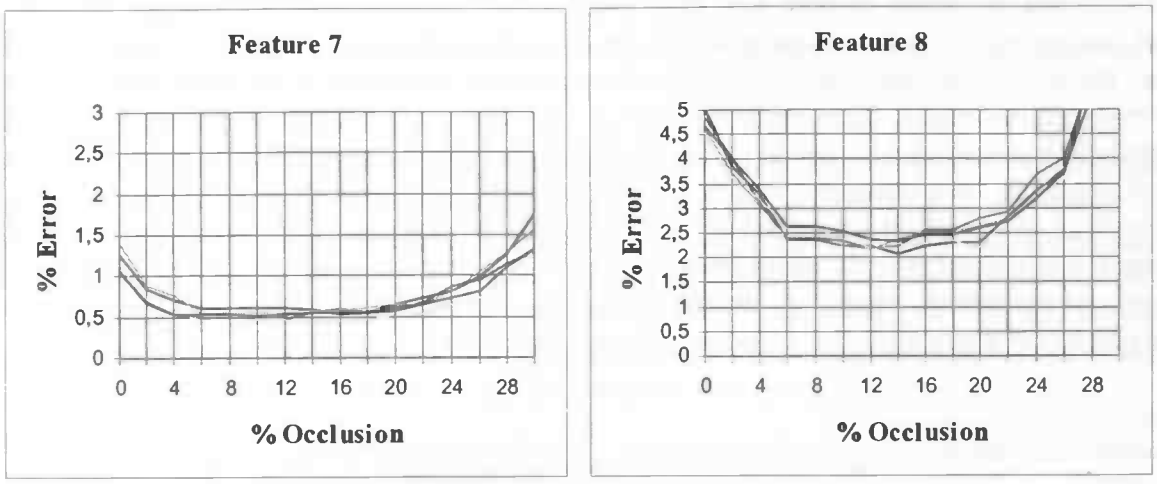


Figure 7-1. By training the neural networks with occluded characters, there is a tremendous improvement in the recognition of occluded characters but also a worsening in the recognition of normal characters. The coloured lines indicated the five different test sets used.

One might expect that now that the neural networks are trained with all percentages of occlusion, equally divided, the recognition rate would be constant over the whole range of occlusion. In other words, the recognition rate of for example 30 percent and 0 percent occluded characters would be the same. After all, the network was trained with as many 30 percent occluded characters as it was with 0 percent occluded ones. For two reasons, this is not the case. First, the more occluded a character is, the less "information" it contains. Although many characters can still be distinguished, the number of characteristics that distinguish them simply decrease. For example, one of the characteristics that distinguish the **R** from the **K** is the closed top of the **R**. This characteristic disappears when the characters are occluded, so there is one less characteristic they can be distinguished on. This is the case with practically all characters. They all contain less "information" if they are occluded. Because of this, we shall see that the error increases if the occlusion increases.

The other effect is that although there were an equal amount of each percentage occluded characters in the train set, in practice, the average occlusion is best trained. The average in this case is 15 percent. The reason for this is that the 0 percent occluded characters have some "support" from a small range of percentages occlusion above 0 percent. The same with 30 percent occluded characters. There is some "support" from a small range below 30 percent. Characters more in the middle of the complete range however have "support" from both sides, both above and below that percentage. In other words, the train set contained more characters of "about" 15 percent occlusion than it did of characters "about" 0 or 30 percent occlusion. This results in a better recognition in the middle of the total range of occlusion than at the edges.

If we combine these two effects, we will see that the recognition is best at about 15 percent and higher at 0 percent than at 30 percent. Figure 7-2 shows this effect and compares the results of the four features. We see that this effect complies to all the features, with feature 7 having the best recognition rate.

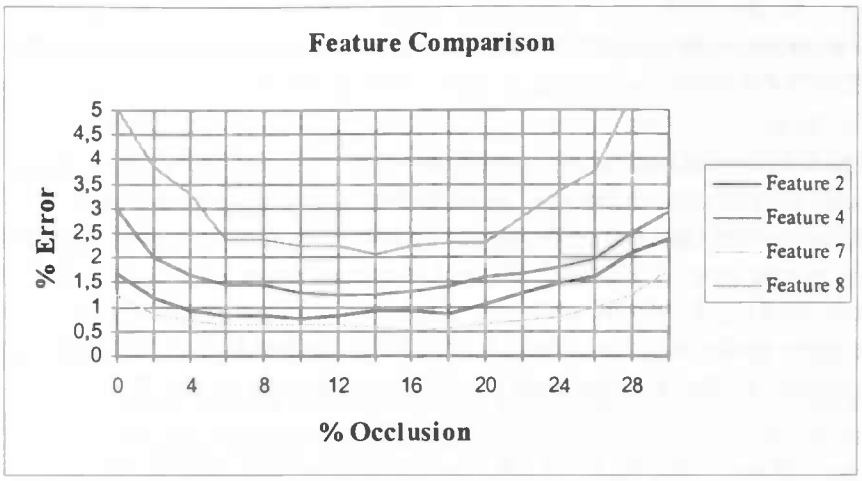


Figure 7-2. Increasing the occlusion has the same effect on all four features with feature 7 having the best recognition rate.

We saw the performances of the four features being trained with all percentages of occlusion. Now we again combine these four features and look at the recognition of the individual classes by the complete OCR. We mentioned that the recognition of the normal characters decreased if we used this new mixed neural network. Figure 7-3 shows how great this loss is. No class solely cause this loss. There is just a general increase in errors that were made. Also the amount of characters that were rejected increased. With not-occluded characters, the percentage correctly recognized characters is 99.0%, with a rejection rate of 0.2% and an error of 0.8%. If we compare this with the results of the traditional neural network, the error has increased with 0.6% and 0.1% more characters were rejected. This means the rate of correctly recognized characters has dropped with 0.7 percent.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400											1												3
2:Z	400							3											1			1		
3	400																							3
4	400			1																				
5:S	400						1								1									
6	400																							2
7	400		1	1						2														
8:B	400						5																	2
9	400				2	3					1													
F	400																	1						1
G	400	1													1									
H	400														1									
J	400	1		2		1				1														
K	290		1				1												2			2		2
L	400																							
N	400																							
P	400																							
R	400												1											1
T	400										1				1						1	1	8	
V	400																1							
X	400																							
Y	212				1			2		1											6			2

8502	2	2	4	3	4	7	5	0	4	2	1	1	0	4	0	1	1	3	0	7	4	8	16

Figure 7-3. If characters with no occlusion are passed through the character recogniser trained for 0% to 30% occlusion, relatively few errors are made.

If we look at the results with 30 percent occlusion (see figure 7-4), we see that the recognition rate has improved drastically compared to the results with our traditional recogniser. Now, with 30 percent occlusion, the percentage correctly recognized characters has increased from 53.3% to 98.3%, the rejection rate from 7.7% to 0.2% and the error from 39.0% to 1.5%. An improvement of respectively 84.4%, 97.4% and 96.2%. Again we cannot say the a particular class is responsible for the error made although the **6** is falsely projected to the **0:D** class in 6 percent of the cases and 3 percent of the **Y** characters are falsely projected to the **T**.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400											1												1
2:Z	400							4																
3	400		1		1			1	1					3							1			2
4	400									5														
5:S	400			2			1																	
6	400	24			1				1															
7	400		2																		1	1		2
8:B	400	9											1											2
9	400							2	1				1								1		1	1
F	400												1					7						
G	400	1																			1			1
H	400								1						1		1							
J	400	1		3						1														4
K	290											1					4		2	2				2
L	400											1												
N	400												1											2
P	400										1													
R	400								1				1		7		2							
T	400										1							1			1		1	1
V	400																1						3	
X	400														2									
Y	212				1															12				1
	8502	35	3	5	3	0	1	7	5	6	2	3	5	3	10	0	8	8	2	14	5	1	5	19

Figure 7-4. If characters with 30% occlusion are passed through the character recogniser, more errors are made than with the 0% occluded characters.

Although the gain with the recognition of occluded characters is enormous, there is a slight loss in the recognition of normal characters. If the occluded characters would become larger in number, this solution could proof useful. For now, we see this loss as unaffordable and have to find a way to leave the recognition of normal characters untouched and still gain performance with the recognition of occluded characters.

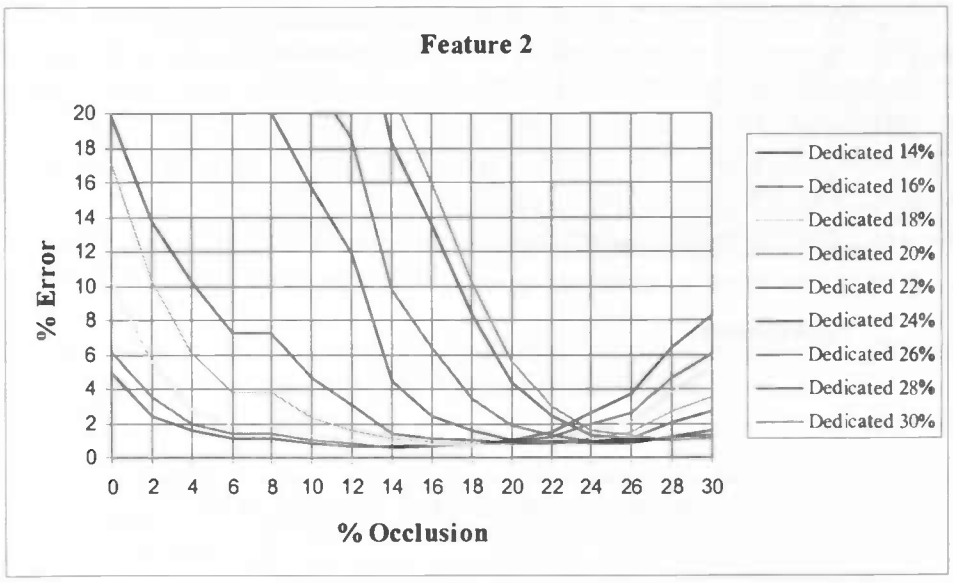
8 Modular neural networks

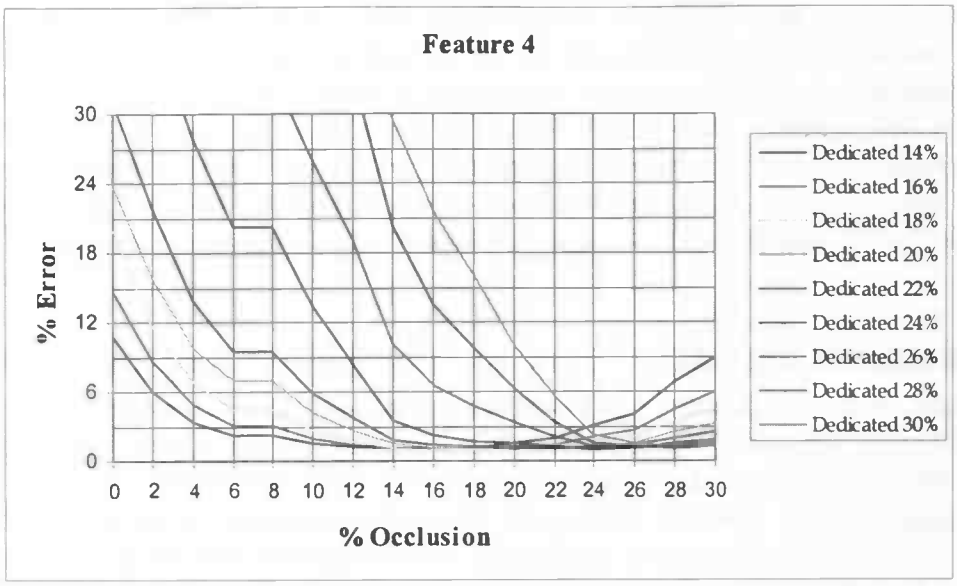
So far, although we have four features, we have used only a single (combined) neural network. In the traditional situation, a neural network that was trained with only not-occluded characters and later a single mixed neural network that was trained with all percentages of

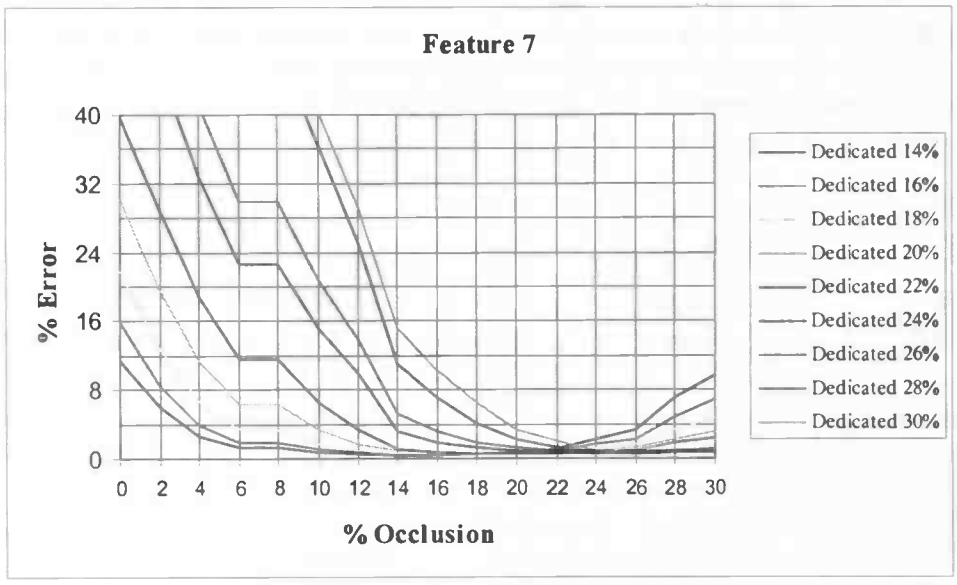
occlusion. In the first case we have a network that performs best with not-occluded characters and performs very bad with occluded characters. In the second case we have a network that does not perform extremely well with a certain percentage of occlusion but in general has a fair performance. In other words, both neural networks have good and bad qualities. By building a modular neural network, thus connecting multiple networks, we can combine those good qualities and achieve better performances at any percentage of occlusion. We will do this by building dedicated neural networks that are trained for a specific percentage of occlusion.

8.1 Dedicated neural networks

With dedicated neural networks, here we mean neural networks that are specifically trained for a certain percentage of occlusion. We expect such neural networks to perform really well for characters with the same amount of occlusion the networks were trained with. Figure 8-1 shows the results with our four features. Here, only the dedicated neural networks for the occlusion range, in our case 14 to 30 percent, are shown. At the occlusion percentage a certain neural network was trained for, it performs best. The more the percentage occlusion of a character differs from the percentage a dedicated network was trained with, the higher the error. If we look at feature 2 for example and look at the 22 percent dedicated network, we see that with 22 percent occluded characters, the error is only 0.8%. The error with 0 percent occluded characters however is as high as 19.8%.




Feature 4
% Error
30
24
18
12
6
0
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
% Occlusion
Dedicated 14%
Dedicated 16%
Dedicated 18%
Dedicated 20%
Dedicated 22%
Dedicated 24%
Dedicated 26%
Dedicated 28%
Dedicated 30%


Feature 7
% Error
40
32
24
16
8
0
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
% Occlusion
Dedicated 14%
Dedicated 16%
Dedicated 18%
Dedicated 20%
Dedicated 22%
Dedicated 24%
Dedicated 26%
Dedicated 28%
Dedicated 30%

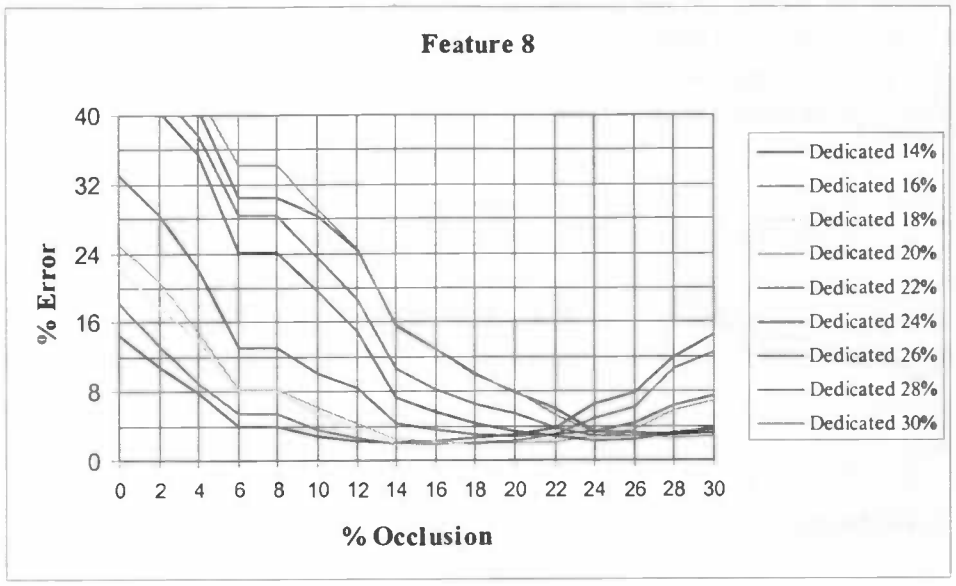
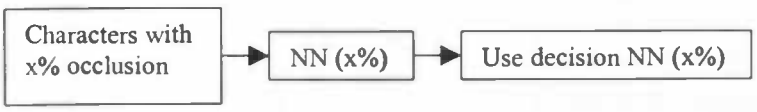


Figure 8-1. Dedicated neural networks have a very good performance with characters with the same amount they were trained with (dedicated to). Performance rapidly decreases for characters with another percentage of occlusion.

We know now that by building a dedicated network, we can achieve a very high recognition rate for the occlusion the network was trained for. If we look at our previous example of a 22 percent dedicated network with feature 2, the error at 22 percent occlusion was 0.8%. If we compare this with the error of our mixed neural network, the error with 22 percent occlusion was 1.1%. An improvement, even though it is small. However, if we compare the performance of a 30 percent dedicated network with our mixed network, the error decreases from 2.2% to 1.1%. This is an even larger improvement. The real improvement however can be found in the dedicated 0 percent neural network, which is in fact our traditional network, in comparison to the mixed network.

The best results can be achieved by building dedicated neural networks and passing characters through the network that fits their amount of occlusion.



The problem here is that there has to be knowledge about the amount of occlusion the characters have before they can be passed through their matching neural network. So far, we have only looked at the recognition rates if we pass a character with a certain amount of occlusion through a certain neural network. Not at any point has there been any knowledge about the supposedly percentage occlusion that the character has. Later we shall see how we can determine this. For now we shall ignore this problem and focus on the gain that can be achieved in theory.

The best performance we can achieve is, as said, if we use a neural network for recognizing a character that matches its percentage occlusion. Figure 8-1 showed us the performance of such a neural network for the whole range of occlusion, in our case 0 to 30 percent. If we combine all the highest recognition rates of those networks, we can see the best recognition

rate that can be achieved in theory for all percentages of occlusion. Figure 8-2 shows the results.

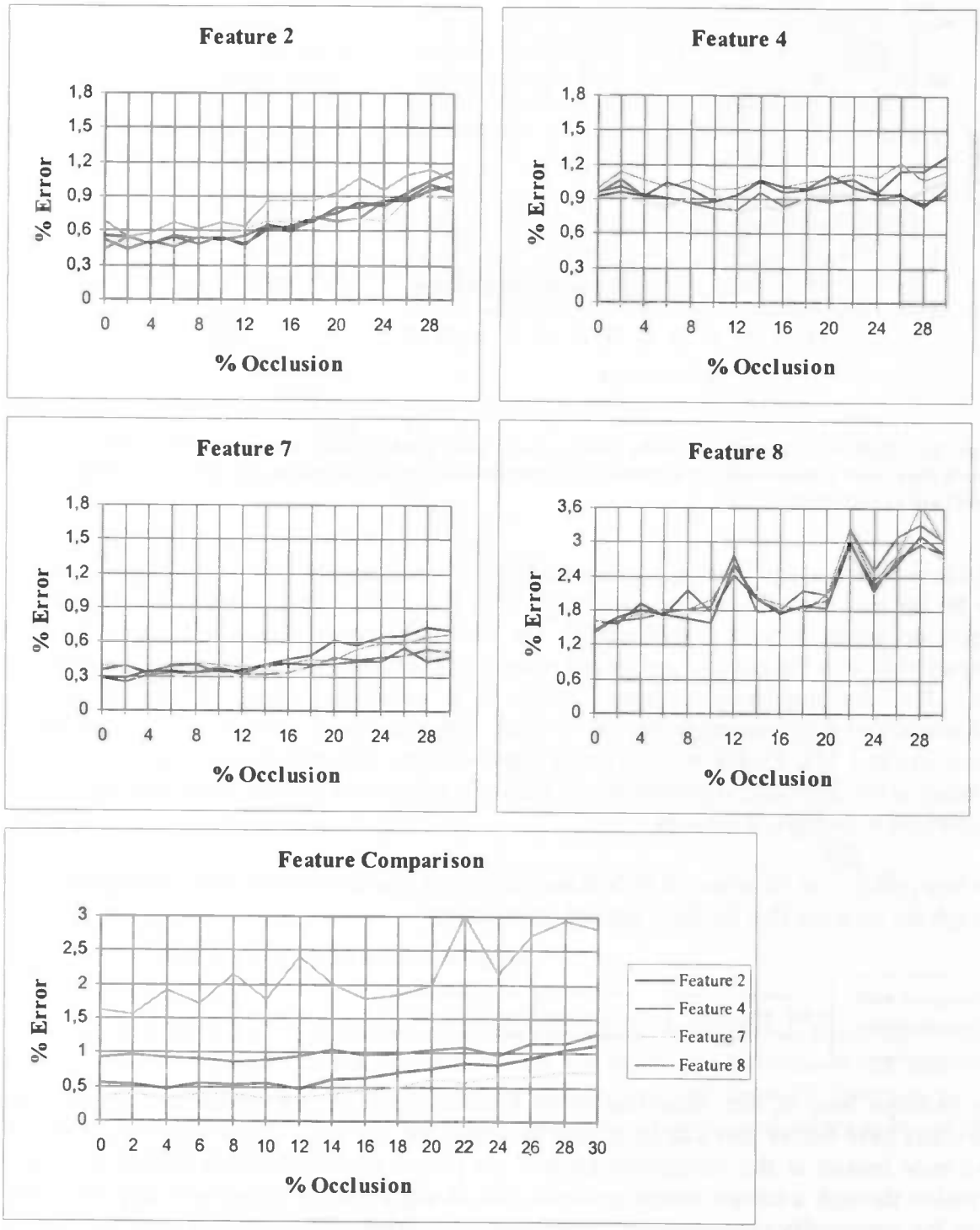


Figure 8-2. If we pass each character through a neural network matching its occlusion percentage, we get the best result that is theoretically possible.

As we have seen with our mixed network, the error slightly increases if we increase the occlusion. This is due to the same reason as explained earlier. If we look at feature 2 and feature 7, it can be seen that this increase starts at about 15 percent occlusion. This nicely shows the effect of the critical point that was explained earlier.

We again combine the four features into our complete OCR and look at the recognition rates of the individual classes. Figures 8-3, 8-4 and 8-5 show the results for a 14, 22 and 30 percent dedicated neural network respectively. Again we see the error slowly increases but no class in particular is solely responsible for this.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400																							1
2:Z	400							3																
3	400																							2
4	400																							
5:S	400						1																	
6	400																							
7	400		2							2														
8:B	400	2																						
9	400					1																		1
F	400																	1						
G	400	1																						
H	400														1									
J	400																							
K	290																			1				1
L	400											1												
N	400																							
P	400																							
R	400	1											1											
T	400		1								1													
V	400																1							
X	400																				1			
Y	212				1															2				
	8502	4	3	0	1	1	1	3	0	2	1	1	1	0	1	0	1	1	0	3	1	0	0	5

Figure 8-3. The OCR dedicated to 14 percent occlusion.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400																							
2:Z	400							4																
3	400					1		1																1
4	400									1														
5:S	400						2																	
6	400	1																						
7	400		2							1												1		1
8:B	400	1											1		1				1					
9	400																							2
F	400																	1						
G	400	1																						
H	400														1				1					
J	400	1		2																				1
K	290															1	1					1		2
L	400											1												
N	400																							
P	400																							

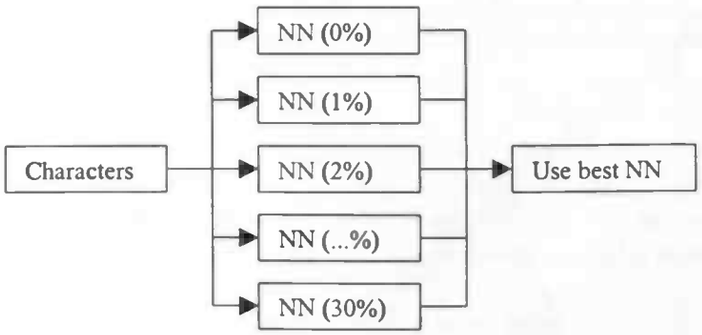
R	400												2				1							
T	400										2													
V	400																1							
X	400																							
Y	212				1															1				2
	8502	4	2	2	1	1	2	5	0	2	2	1	3	0	2	1	3	1	2	1	0	2	0	9

Figure 8-4. The OCR dedicated to 22 percent occlusion.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400																							1
2:Z	400							3																
3	400								1															
4	400									1														
5:S	400						2																	
6	400								1															1
7	400		2							1				1							1			
8:B	400						7												1					
9	400					1																		1
F	400												1					3						
G	400	1					1																	
H	400								1						1									
J	400	1		1		1																		1
K	290												1				1			1				1
L	400											1												
N	400												1											
P	400										2													
R	400												2											1
T	400		1								1							1						
V	400																2							
X	400																							
Y	212				1															3				2
	8502	2	3	1	1	2	10	3	3	2	3	1	5	1	1	0	3	4	1	4	1	0	0	8

Figure 8-5. The OCR dedicated to 30 percent occlusion.

The ideal situation would be where the characters are passed through their matching dedicated neural network. As said, for the moment we do not determine beforehand the amount of occlusion of a character, so we cannot match it with its dedicated network. One way to solve this is to pass each character through every dedicated network and afterwards somehow decide which network to believe.



This could for example be done by looking at the neurons. The network with the highest output or largest difference with the other neurons could be the winner. Also a decision could be made based on the results of all networks. The class that has the most wins is decided the winner.

This method could achieve a very good performance but every character has to be passed though all dedicated networks. This is of course a very brute force solution. Mostly because of the fact that most dedicated networks are very seldom the best choice. In far most cases the neural network that was trained with not-occluded characters will be the winner. Perhaps in some cases the 1 or 2 percent dedicated network scores best if a character was not occluded, but in no cases would the 0 percent dedicated network not have won if the 1 and 2 percent dedicated ones were left out. In other words, this method is not only brute force but also has some overkill. This is why we shall now make a selection of which networks to use in our modular network. We shall see that by using only two we can achieve similar results.

8.2 Ranged neural networks

Previously, when we spoke of occlusion of license plates, we concluded that the amount of occlusion was ranged between about 15 and 30 percent. Later we concluded that the lower bound here, 15 percent, was also a critical point. With this knowledge we can simplify our previously described large modular network into a two component modular network. With this modular network we again divided the characters into two groups. Characters are either normal or occluded. For the normal or not-occluded characters we have our traditional 0 percent dedicated network. For the occluded characters we shall use a new ranged network, which is trained with occluded characters ranging from 14 to 30 percent.

We saw that our mixed neural network, which was trained with all percentages of occluded characters ranging from 0 to 30 percent, fairly good recognition results were achieved in the occlusion area (15 to 30 percent). Now that we use our traditional network for the not-occluded characters, the network dedicated to occlusion does not have to be trained with these normal characters. For the occlusion area we now build a neural network that was trained with occluded characters ranging from 14 to 30 percent. We shall see later that the recognition rates of this network are comparable to the rates achieved by the dedicated networks.

First we shall look at the recognition rates of the individual four features. Figure 8-6 shows the results. We see that the performance in the occlusion area the networks were trained with, is very good. Of course below this range, the error increases rapidly. Again, all features behave in the same way with feature 7 having the best results. Notice that feature 2 is the less sensitive to occlusion below 15 percent.

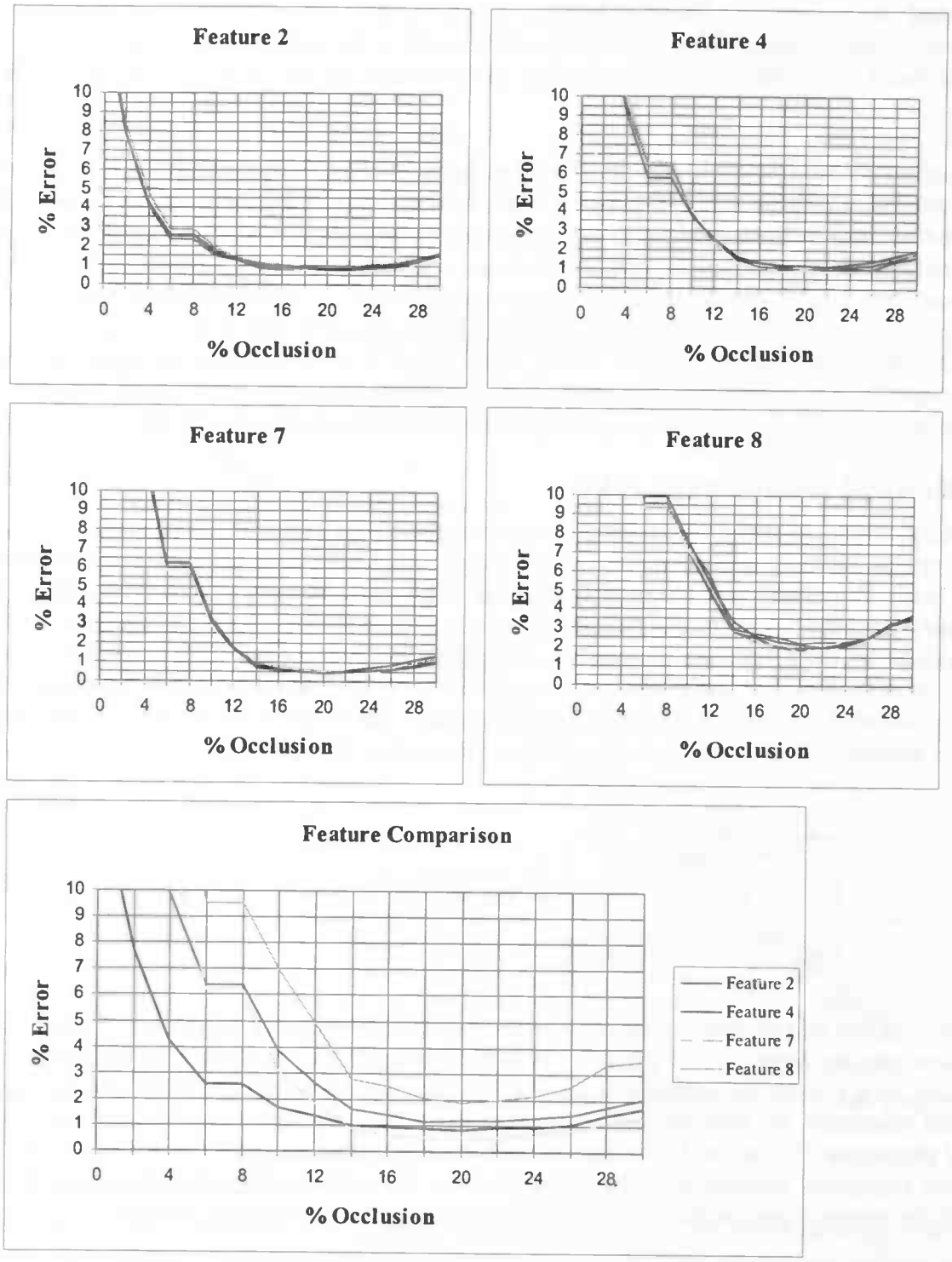


Figure 8-6. The performances of a ranged network are very good within its range but rapidly decrease outside of it. The performance of this network within its range is comparable to the performances of the individual dedicated networks.

If we combine our four features again, we can look at the performance of the complete ranged network. As we saw with the individual features, our ranged network performs very badly with not-occluded characters. Figure 8-7 shows the results. However this is not a problem because the other component in our modular neural network handles these characters.

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?

0:D	400						204		1										1				2
2:Z	400					6		3															5
3	400					39																	8
4	400		2	8																			1
5:S	400			7			4		1														6
6	400		15	24																			4
7	400				1	1				10								1				1	1	2
8:B	400			1	1	1	201																2
9	400	2			25	1	6				1												39
F	400							3							2			8					8
G	400					3	109		14										6				34
H	400				1										1								4
J	400	1		2																			1
K	290		1		1											1			113				5
L	400																						
N	400												20						3				3
P	400						1				115												1
R	400				2		2				3		1		10		1						12
T	400										1				1							1	278	3
V	400	2																				1		1
X	400					1													23					14
Y	212				2			2		37								1				73		6
	8502	5	18	42	33	52	527	8	16	47	120	0	21	0	14	1	1	10	146	0	0	76	279	161

Figure 8-7. Below its lower bound of 15 percent, the error of the ranged network increases rapidly. Here we see the recognition of not-occluded characters.

The recognition rate of the ranged network with for example 22 percent (average occlusion in practice) and 30 percent (our maximum percentage) occluded characters is very high. In fact, as we shall see later on, they are comparable to the recognition rates of our dedicated networks. With 22 percent occluded characters, the percentage correctly recognized characters is 99.4%, with a rejection rate of 0.1% and an error of 0.5%. With 30 percent occluded characters, this is 99.3%, 0.1% and 0.6% respectively (see figure 8-8 and 8-9).

	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400												1											1
2:Z	400							3																
3	400					1			1															1
4	400																							1
5:S	400						2																	
6	400	1			1																			
7	400		2							1														1
8:B	400	1											2											1
9	400								1															
F	400																	1						1
G	400	1																						
H	400								1						1									
J	400	1		1																				1
K	290	1														1				1		1		1
L	400											1												
N	400																							

P	400										1													
R	400												2											1
T	400										2												1	2
V	400																1							
X	400														1									
Y	212				1																1	1		
	8502	5	2	1	2	1	2	3	3	1	3	1	5	0	2	1	1	1	0	1	1	2	1	11

Figure 8-8. With 22% occluded characters, the ranged network has an error of only 0.4% and a rejection rate of 0.1%.

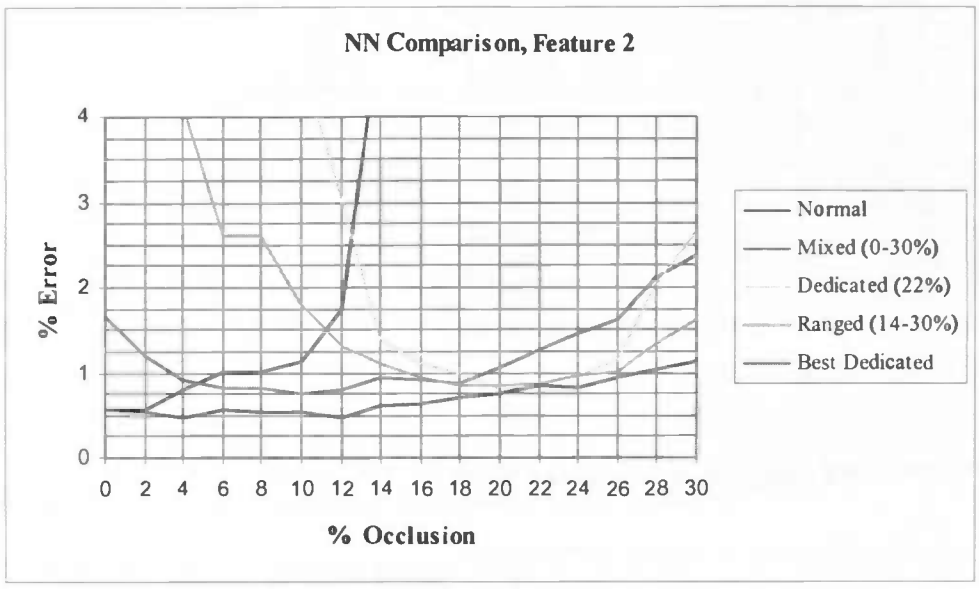
	Total	0:D	2:Z	3	4	5:S	6	7	8:B	9	F	G	H	J	K	L	N	P	R	T	V	X	Y	?
0:D	400																		1					
2:Z	400							4																1
3	400					1			1					2							1			1
4	400									2													1	1
5:S	400						2																	
6	400	10			1				1															
7	400		2																		1			1
8:B	400	2					7						1								1			1
9	400					1																		1
F	400																	8						
G	400	1																						
H	400								1						1									
J	400	1																						1
K	290															2	1							1
L	400											1												
N	400																							
P	400										2												1	
R	400								1						3		1							
T	400										1			1							1			2
V	400				1												1						1	1
X	400														4									1
Y	212				1															12				
	8502	14	2	0	3	2	9	4	4	2	3	1	1	3	8	2	3	8	1	12	4	0	3	12

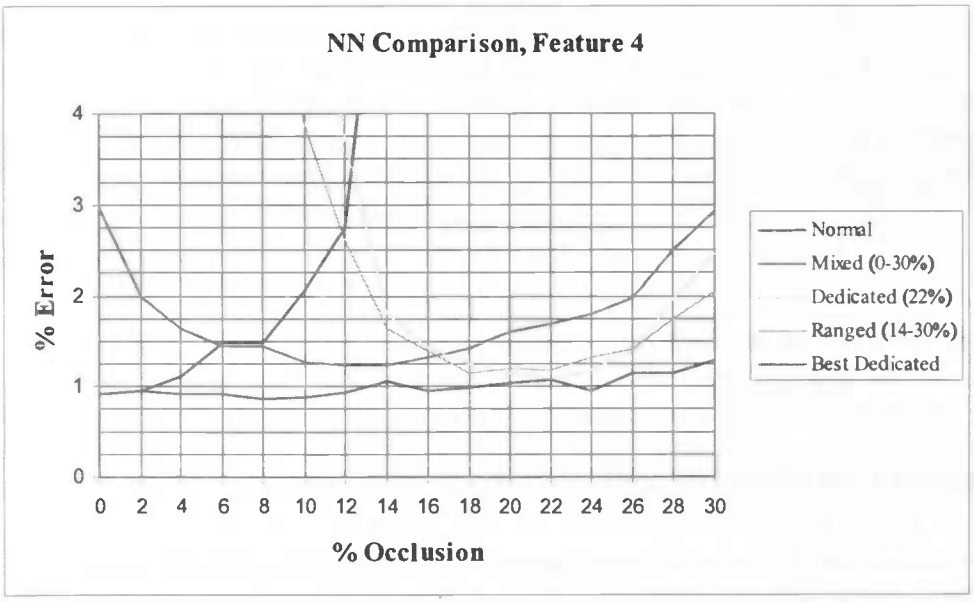
Figure 8-9. With 30% occluded characters, the ranged network has an error of 0.6% and a rejection rate of 0.1%.

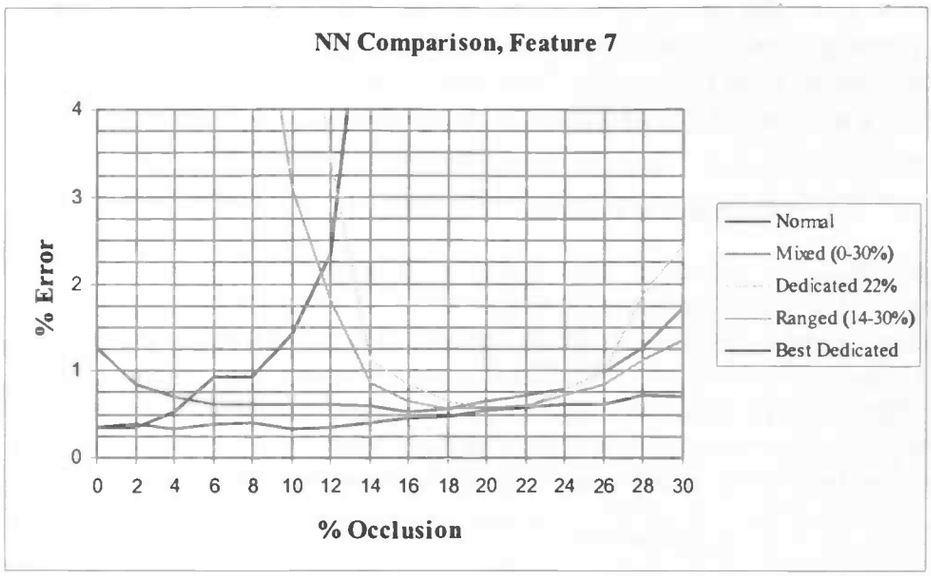
8.3 Comparison

So far, we have discussed several types of neural networks. Our initial network was trained with not-occluded characters only. Then we build a network that was trained with all amounts of occlusion from 0 to 30 percent. Dedicated networks were built to recognize a particular percentage of occlusion. Finally a ranged network was trained with occlusion percentages ranging from 14 to 30 percent. Here we will give a comparison of the recognition rates of all these different neural networks. Figure 8-10 shows these rates for all four features. The graphs nicely show that in the ideal situation, every character is passed through a neural network that matches its occlusion. The recognition rates that can be achieved then are shown by the "Best Dedicated" lines.

If we look at characters as either normal or occluded, we see that by using the normal network for not occluded characters and the ranged network for occluded characters (15% and more), we can achieve recognition rates close to the "best dedicated" results. This is what we were preparing for when we constructed this ranged network previously.







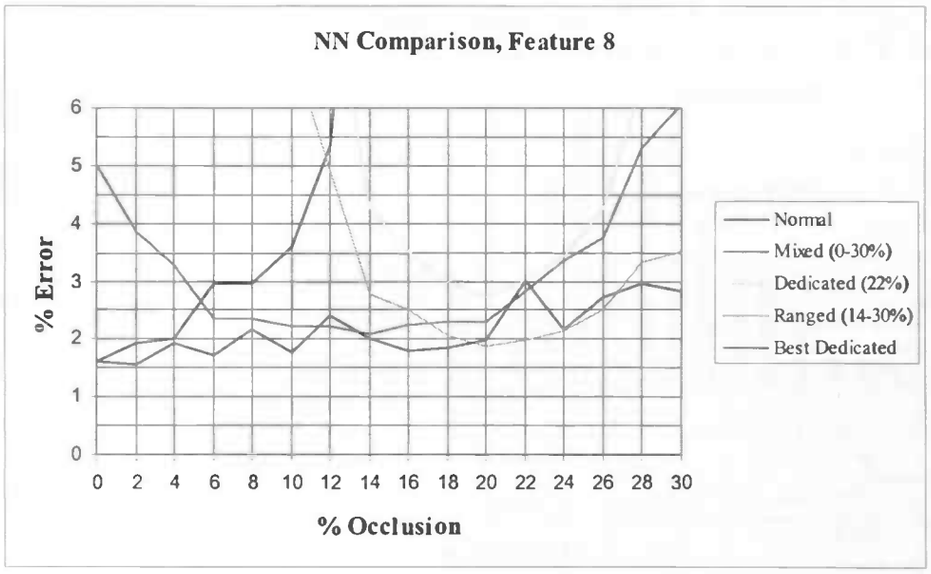


Figure 8-10. A comparison of the different types of neural networks described.

Figure 8-11 again shows the different recognition rates, but now for when the four features are combined. The results with the features showed that by using the ranged network, performances could come close to the ones achieved using dedicated networks. Again we see that this is very much the case. We took 22 and 30 percent as our example. If we compare the recognition rates of the dedicated networks and the ranged network for these percentages, we see that the dedicated networks perform only slightly better. The rejection rates are the same but the dedicated networks have an error of 0.4% and 0.6% compared to 0.5% and 1.0% respectively. For this reason, we shall build a modular neural network, which contains both the traditional network and our new ranged network. Previously when we talked about such a modular network, we passed every character through each network. Based on the results of, in this case, both networks, a final decision had to be made. Even though we simplified our previous approach of passing each character through many dedicated networks, while far most characters are not occluded, the ranged network is unnecessarily used in most cases. For this

reason we will look at ways to detect occlusion so our ranged network will only be used in case of occluded characters.

Train set	Test set	% Correct	% Reject	% Error
0% (traditional)	0%	99.7	0.1	0.2
0% (traditional)	22%	81.8	2.8	15.4
0% (traditional)	30%	53.3	7.7	39.0
0%-30% (mixed)	0%	99.0	0.2	0.8
0%-30% (mixed)	30%	98.3	0.2	1.5
22% (dedicated)	22%	99.5	0.1	0.4
30% (dedicated)	30%	99.3	0.1	0.6
14%-30% (ranged)	22%	99.4	0.1	0.5
14%-30% (ranged)	30%	98.9	0.1	1.0

Figure 8-11. A comparison of the recognition rates of our different neural network shows that with occluded characters, a huge gain in performance can be achieved.

9 Occlusion detection

If we do not want to use more neural networks than necessary, it is important to somehow detect whether a character is occluded. In other words we will need some sort of classifier. Since we are focused here on the occlusion of license plates, we will take one step back and do not try to detect occlusion for a single character but look at the license plate as a whole. This means we detect occlusion for six characters at a time. We shall do this first by looking at the width / height ratio of the license plate.

9.1 Width / height ratio

Occlusion in our case causes a license plate to lose a part of its top. This causes the height of the license plate to decrease. The width on the other hand remains the same. In other words, the relation between the width and height changes. In theory, this relation between width and height is fixed for a certain type of license plate. The value of this relation in theory is not very important. Of more importance is the value found after the license plate is located in the total image and cut out. In figure 9-2 we show the width and height in pixels of about a thousand segmented license plates. The range in this case is explained by the fact that the vehicles are not at a fixed distance when the image is taken. Of course, the further the vehicle is away, the smaller the width and height of the license plate. Of importance here is the relationship between those two values. In theory, this should be a fixed value. Figure 9-3 shows us that this is not the case. The peek at the beginning of the x-axis is caused by rectangle license plates, most used on motorcycles. The second peek shows the entire range of the width / height relation of the regular plates. Notice that the less high a plate is to its width, the higher this value. This means that license plates that are occluded have a high value for this relation.

With a set of a thousand not-occluded license plates, the width / height relation varied from about 3.3 to 5.2. If we want to decide whether a license plate is occluded or not, we must set a threshold. If the value is higher than this threshold, the plate is considered to be occluded. Since not-occluded license plates are the far largest group, we must set this threshold high. This way we limit the chances of treating a not-occluded license plate as being occluded.

We will set the threshold at 5.2. If we take a width of 69 and a height of 16 pixels, the width / height relation becomes 4.3. This would be a common situation. What happens if we cut of 20 percent of the license plate? The height becomes 13 (16 – 20%). Of course the width remains the same. Now the width / height relation is 5.3 (69 / 13). Just above our threshold of 5.2. In this case, the occluded license plate is correctly treated as an occluded one.

However, in many cases the license plate is only 15 percent occluded. If we cut of 15 percent from a height of 16 pixels, 14 pixels remain. The width / height relation now becomes 4.9 (69 / 14). In this case, the occluded license plate is mistakenly treated as a normal plate. This problem occurs even sooner if a lower value for 4.3 is chosen. Of course for a higher value of for example 4.6, the problem no longer occurs, but lower average width / height values are just to large in number. By setting the threshold to a lower value, the problem is not solved either. By doing so, many not-occluded license plates are treated as occluded ones which results in not recognizing them.

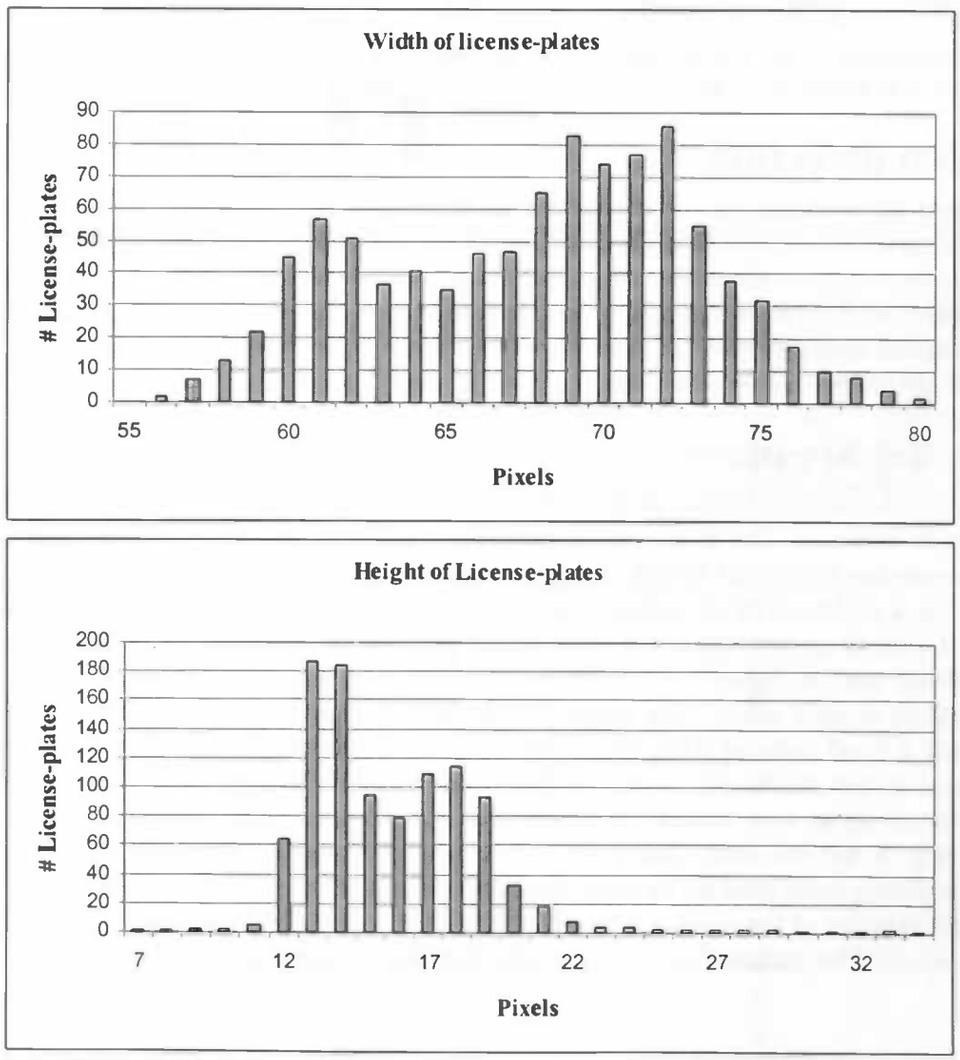


Figure 9-1. Because images of license plates are taken from ranging distances, the width and height of a license plate in pixels varies.

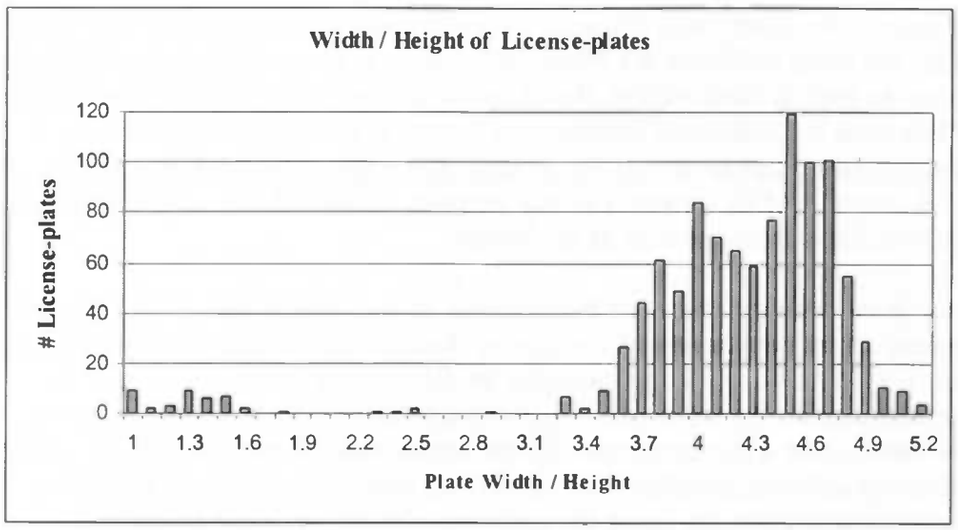
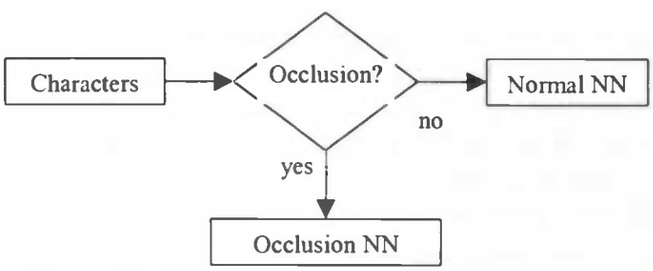


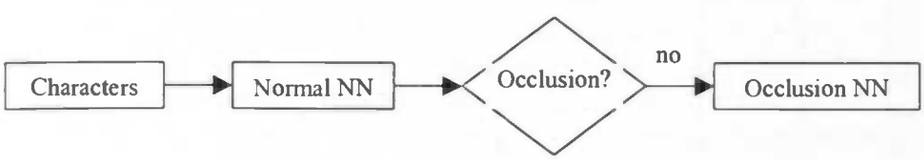
Figure 9-2. By looking at the width / height relation of a license plate, it proves difficult to make a decision whether the license plate is occluded or not.

10 Occlusion classifier

By looking at the relationship between the width and height of a license plate, occlusion can, although not in all cases, be detected before neural networks are used to recognize the characters. This method is shown below.



We shall now see that we can also use the output of a neural network to determine whether there was occlusion. In this case we first pass characters through the traditional neural network and determine afterwards if the characters were occluded. If so, they are passed through a neural network trained for occluded characters. The decision whether to use this second network will be based on the confidence levels given to the characters by the first network.



10.1 Confidence levels and occlusion

Throughout this article we often spoke of the four features and combining them into the complete OCR. The four features each have 22 output neurons, one for each class. When we said the character was correctly recognized, we meant that the neuron belonging to that class

had the highest output. No notice was given to the runner up, or the class with the second highest score. Also the score itself was not important, as long as it was highest. By combining these networks, we do look at those values. Based on all four networks, a final winning class is computed. While with the individual features, we looked at the output neurons, now a so-called confidence level is given to the winning class. The higher this confidence lever, the more certain the network is of its answer. For our purpose, seven different confidence levels are given with 0.99 as the highest and 0.65 as the lowest.

So the confidence level is an indication of the certainty of the correctness of the recognized character. If we pass a very good image of a character through the recogniser, in most cases a confidence level of 0.99 is given. If the character is distorted, a lower confidence level is given. If the character cannot be recognized, it is rejected with the "?", which has of course no confidence level. Notice that a character can also be falsely recognized with any confidence level. This relationship between distortion and confidence level also applies to occlusion. On average, the greater the occlusion, the lower the confidence level (see figure 10-1).

0.99 0.84 ?

Figure 10-1. There is a relationship between the occlusion and confidence level of a character.

However, this relation is not secure enough to be helpful. Although on average, there is a very nice relation between these two values, in practice we deal with individual characters, not averages. For example, the average confidence level given to a 22 percent occluded **H** is 0.91. However, in little over 50 percent of the cases, the 22 percent occluded **H** is given a confidence level of 0.99. This problem occurs with every class as is nicely shown in figure 10-2 and 10-3. So by looking at the confidence level of a single character, little can be said about the occlusion it may have.

	Total	0.99	0.96	0.87	0.84	0.77	0.74	0.65
0:D	401	380	8	1	7	3	1	1
2:Z	400	375	14	1	4	3	2	1
3	400	377	4	2	11	1	2	3
4	401	386	5	0	6	2	1	1
5:S	399	383	6	2	4	2	1	1
6	401	387	4	1	4	2	3	0
7	400	364	13	3	15	2	2	1
8:B	399	356	8	6	21	2	3	3
9	400	385	7	1	3	1	1	2
F	401	383	6	1	7	1	3	0
G	400	382	8	0	7	0	0	3
H	400	376	11	2	7	2	2	0
J	397	383	6	4	3	0	1	0
K	288	267	6	3	5	1	3	3
L	401	391	3	2	3	2	0	0
N	402	384	9	0	7	1	0	1
P	400	390	3	0	5	2	0	0
R	397	377	4	2	6	3	3	2

T	399	382	8	0	5	1	3	0
V	399	390	0	0	5	1	2	1
X	400	383	5	2	9	0	1	0
Y	211	197	2	1	4	3	3	1

Figure 10-2. The distribution of the confidence levels given to not-occluded characters by the normal neural network.

	Total	0.99	0.96	0.87	0.84	0.77	0.74	0.65
0:D	401	160	61	21	94	68	106	103
2:Z	400	159	46	31	61	45	35	14
3	400	26	25	13	53	75	82	50
4	401	97	31	40	69	46	38	41
5:S	399	68	50	26	129	46	33	28
6	401	40	16	20	90	11	48	27
7	400	17	80	95	88	61	43	45
8:B	399	29	8	13	63	73	73	79
9	400	19	6	6	46	16	87	89
F	401	32	9	15	57	52	62	62
G	400	95	32	21	70	41	43	45
H	400	218	45	4	57	54	14	17
J	397	385	7	2	12	27	34	9
K	288	144	30	52	149	45	118	71
L	401	392	3	1	6	0	2	21
N	402	270	65	17	23	17	15	24
P	400	39	18	27	88	70	112	63
R	397	2	0	1	2	8	11	17
T	399	190	24	19	49	40	48	112
V	399	94	34	10	59	68	57	81
X	400	52	22	49	92	21	88	44
Y	211	5	7	4	19	15	85	77

Figure 10-3. The distribution of the confidence levels given to 22% occluded characters by the normal neural network.

By looking solely at one character, there is no usable relation between its confidence level and its occlusion. If we again look at figure 10-2 and 10-3, we see that with no occlusion, most characters have a confidence level of 0.99 and only a few have lower ones. With 22 percent occlusion, still most characters have a confidence level of 0.99, but now many more have lower ones. So the chance for a 22 percent occluded character to have a confidence level of 0.99 is far smaller than for a not-occluded character. This effect will prove useful. We will divide the confidence levels into two groups. A character either has a confidence level of 0.99 or a lower one. So we combine all confidence levels other than 0.99 into one group. Figure 10-4 shows this relationship between occlusion and the chance for a character to have a confidence level of 0.99. We shall use this division to determine whether characters have occlusion or not. Of course we cannot look at one character and decide whether it is occluded or not based on whether is has a confidence level of 0.99 or not. However, if we look at more characters at a time, this decision can indeed be made. Remember that in our case, if a license plate is occluded, all six characters are equally occluded. We shall see that by looking at the confidence levels of those six characters and more specifically at the occurrences of confidence levels of 0.99, we can statistically very well determine if a license plate and hence its characters were occluded.

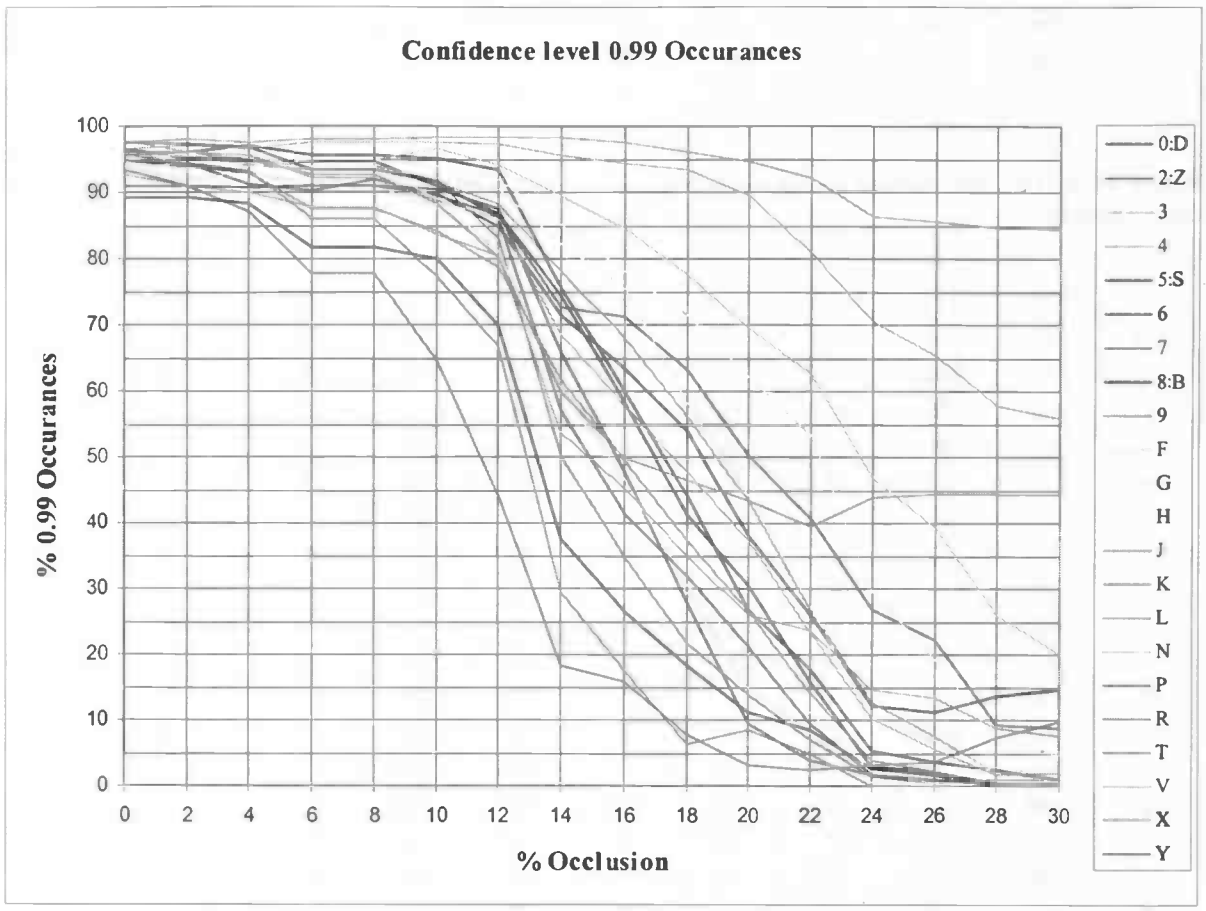


Figure 10-4. Relationship between occlusion and confidence level 0.99.

10.2 Confidence levels as a classifier

The chance for a random not-occluded character to have a confidence level of 0.99 is 95.1%. If we look at six characters at a time that are all equally occluded, the chance that four or more of those characters have a confidence level of 0.99 is 99.8% (see figure 10-5). So if four out of the six characters have a confidence level of 0.99, we are right in 99.8% of the cases if we say the license plate is occluded. If a license plate was for example 22 percent occluded, there is a 92.6% chance that three or more characters have a confidence level lower than 0.99. In case of 30 percent occlusion, this chance is even higher with 99.0%. We shall use this threshold to determine whether the characters were occluded or not. If based on this threshold, the characters are supposedly occluded, they are passed though the ranged network, which was described earlier.

Not-occluded characters	
Times confidence level of 0.99	Chance
6	74.0%
5	22.9%
4	2.9%
3	0.2%
2	0.0%
1	0.0%
0	0.0%

22 percent occluded characters	
Times confidence level of 0.99	Chance
6	0.0%
5	0.1%
4	6.3%
3	19.2%
2	32.6%
1	29.6%
0	11.2%

Figure 10-5. By setting the threshold to 4, only 0.2% of the not-occluded characters are falsely treated as occluded ones. With 22 percent occluded characters, this error is 7.4%.

Occlusion?
If #confidence levels of 0.99 4 then characters are passed through the ranged network. If not, we are done.

However, in 0.2% of the cases we mistakenly treat a not-occluded license plate as an occluded one. Also in 7.4% a 22 percent occluded license plate is treated as a not-occluded one. This error can be reduced by doing the occlusion check again after we have used the ranged network. This is also very logical. After having the characters passed though the ranged network, we want better confidence levels than the ones that were given by the normal network. We can use the same check. Again we want four or more characters to have a confidence level of 0.99. If this is now the case, it is decided the characters were indeed occluded. If this is again not the case, it is decided that the low confidence levels that were given the first time were probably not due to occlusion. In this case the output of the ranged network is ignored and the output of the first network is final.

Occlusion? (Second time)
If #confidence levels of 0.99 4 then it is assumed the characters were indeed occluded. If not, it is assumed the characters were wrongfully treated as occluded.

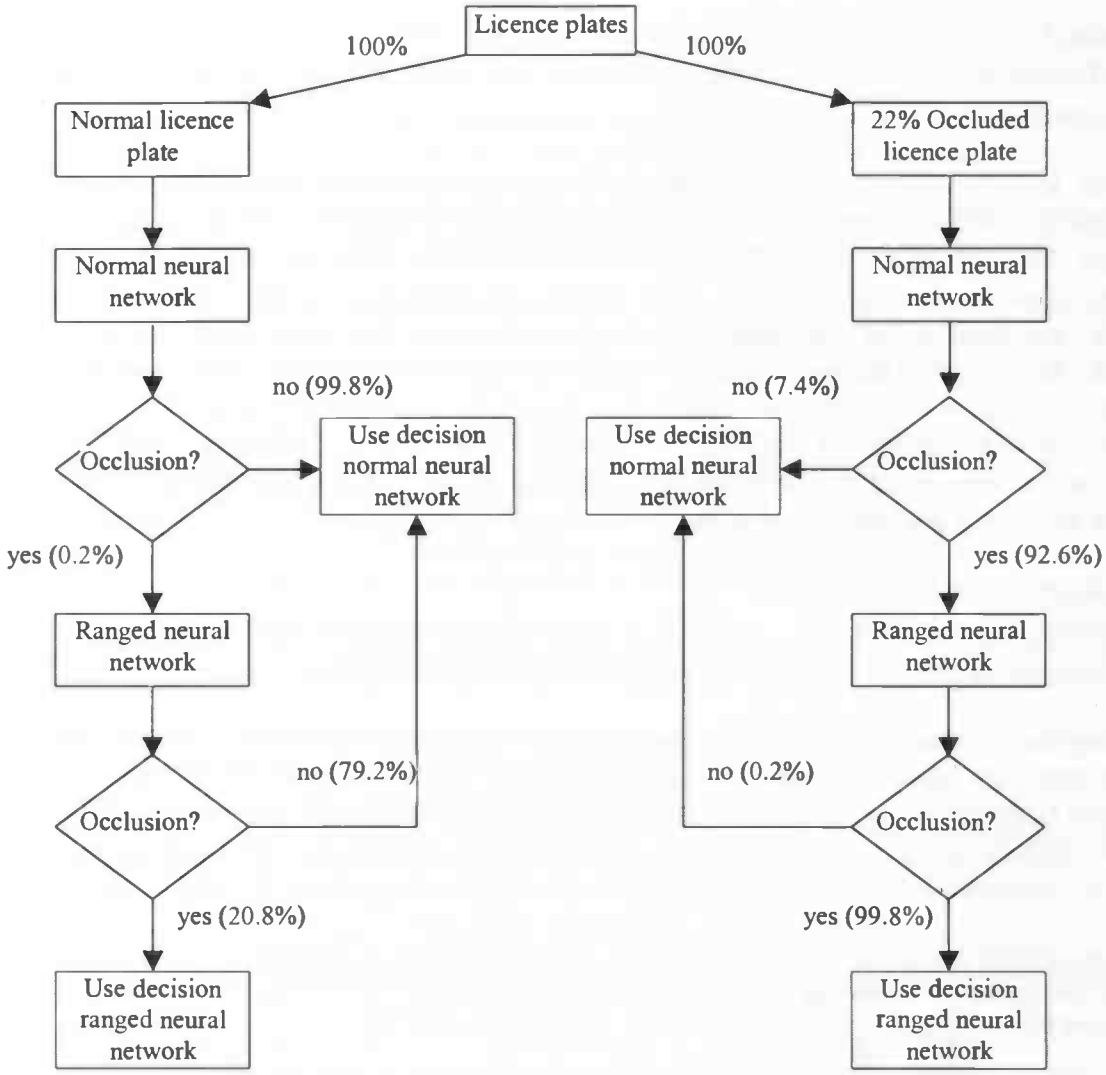
By doing this second check, a large percentage of the small error made with not occluded license plates is being corrected. Of the 0.2% error made here, 79.2% is corrected. The downside however is that, in case of 22 percent occluded characters, also 0.2% of the 96.6% is wrongfully corrected (see figure 10-6). Because of the fact the not occluded characters are far larger in number, this loss is more than compensated by the gain in the other case.

Not-occluded characters	
Times confidence level of 0.99	**Chance**
6	0.1%
5	13.5%
4	7.2%
3	20.5%
2	32.8%
1	28.0%
0	9.9%

22 percent occluded characters	
Times confidence level of 0.99	**Chance**
6	73.0%
5	23.6%
4	3.2%
3	0.2%
2	0.0%
1	0.0%
0	0.0%

Figure 10-6. By applying a second check, a large percentage of the first error is corrected.

The figure below shows how this is done. It shows what happens if a plate is not occluded and if it is 22 percent occluded.



As far most characters are not occluded, the ranged network will (with a small error) only be used if the characters were occluded. This way, no time is wasted. What effect does this method have on the recognition of occluded characters and more importantly, does this effect the recognition of not occluded characters. Also because of the feedback created by the second confidence level check, the recognition rates of not-occluded characters remain unchanged. The recognition rate of 22 percent occluded characters (the most common) increases from 81.8% to 98.0% (see figure 10-7). A great improvement.

Occlusion	Network	% Correct	% Reject	% Error
0	Normal	99.7	0.1	0.2
22	Normal	81.8	2.8	15.4
0	Confidence level Classifier NN	99.7	0.1	0.2
22	Confidence level Classifier NN	98.0	0.4	1.6

Figure 10-7. The recognition rate for 22 percent occluded characters increases from 81.8% to 98.0% while leaving the rates of not-occluded characters untouched.

11 Conclusion and summary

The performance of an ALPR system or Automatic License Plate Recognition system significantly decreases if the license plate is occluded. Recognition rates of characters of such a license plate can drop 20 to 50 percent if the characters are respectively 20 or 30 percent occluded. The problem lies mainly in the OCR. The neural networks that were trained with not-occluded characters make more errors if the occlusion increases. An occlusion of 15 percent proves to be a critical point. Above this percentage, the error increases more rapidly. By including occluded characters in the train set of the neural network, the recognition of occluded characters increases greatly. The unacceptable downside is that with this neural network, the recognition of not-occluded characters decreases slightly. Since occluded characters are rare, the gain in recognizing them cannot compensate this loss. The ideal situation would be to have each character be recognized by a dedicated network that matches its occlusion. Such a dedicated network was only trained with one percentage of occlusion. The problem here is that a way must be found to detect the amount of occlusion before passing the character through its matching neural network. Another solution would be to pass each character through every dedicated neural network and determine afterwards which network did best. A good solution proved to be to see characters as either not occluded or occluded. The occlusion ranges from 15 to 30 percent, since this proved to be the case with license plates in practice. To recognize these occluded characters, we use a neural network that was trained with occluded characters ranging from 15 to 30 percent, equally divided. Since not-occluded license plates are far greater in number than occluded ones (for example 100 to 1), we will first pass all six characters through the normal traditional neural network. Then, based on the confidence levels given to the six characters, there is decided whether the characters are probably occluded. This is done because there proved to be a relation between the confidence levels that were given to the characters and their occlusion. If it was decided that the characters were probable occluded, they are passed though the second network, which was specifically trained with occluded characters. If not, no further adjustments are made. By doing another confidence level check after the characters are passed through the second network, the small percentage error that was made with the first decision is partially corrected. With this method, the recognition rate of not-occluded character remained the same but with 22 percent occluded characters (the average in practice with license plates), the recognition rate increases from 81% to 98%.

12 References

[1] J. Bulas-Cruz, J. Barroso, A. Rafael and E.L. Dagless, "Real-Time Number Plate Reading".

[2] C. Busch, R. Dörner, C. Freytag and H. Ziegler, "VESUV – a Computer Vision System for Integrated Traffic Analysis", in: *Proceedings of the 1998 International Conference on Image Science, Systems and Technology (CISST'98)*, Las Vegas, pp. 310-316, 6-9 July, 1998.

[3] P. Davies, N. Emmott and N. Ayland, "License Plate Recognition Technology for Toll Violation Enforcement", in: *Proceedings IEE Colloquium on Image Analysis for Transport Applications*, London, UK, pp. 7/1-7/5, 16 February, 1990.

[4] S. Draghici, "A Neural Network based Artificial Vision System for License Plate Recognition", *International Journal of Neural Systems*, Vol. 8, No. 1, pp. 113-126, February 1997.

[5] W. Feng, Y.Y. Yan, G. Jin, M. Wu and Q. He, "Optoelectronic Hybrid License Plate Recognition System", in: *Proceedings of the SPIE (Application of Digital Image Processing XXII, Denver, Colorado, 20-23 July, 1999)*, Vol. 3808, pp. 685-692, 1999.

[6] T. Hori, "Traffic camera system development", in: *Proceedings of the SPIE (Real-Time Imaging II, San Jose, California, 10 February, 1997)*, Vol. 3028, pp. 81-90, 1997.

[7] C.M. Hwang, S.Y.Shu, W.Y. Chen, Y.W. Chen and K.P. Wen, "A PC-based Car License Plate Reader", in: *Proceedings of the SPIE (Machine Vision Applications, Architectures and Systems Integration, Boston, Massachusetts, 17-18 November, 1992)*, Vol. 1823, pp. 272-283, 1992.

[8] K. Kanayama, Y. Fujikawa, K. Fujimoto and M. Horino, "Development of Vehicle-License

Number Recognition System Using Real-Time Image Processing and its Application to Travel-Time Measurement", in: *Proceedings of the 41st IEEE Vehicular Technology Conference*, pp. 708-804, 19-22 May, 1991.

[9] H.C. Lee and C.S. Jong, "A Hybrid Neuro-Fuzzy Approach for Automatic Vehicle License Plate Recognition", in: *Proceedings of the SPIE Conference on Applications and Science of Computational Intelligence*, Vol. 3390, pp. 159-168, April 13-16, 1998.

[10] T. Naito, T. Tsukada and K.I. Yamada, "Moving-Vehicle License Plate Recognition Method Robust to Changes in Lightning Conditions", *Systems and Computers in Japan*, Vol. 31, No. 11, pp. 82-91, 2000.

[11] T. Naito, T. Tsukada, K. Yamada, K. Kozuka and S. Yamamoto, "Robust Recognition Methods for Inclined License Plates under Various Illumination Conditions Outdoors", in: Proceedings of 1999 Intelligent Transportation Systems Conference, Tokyó, Japan, pp. 697-702, 5-8 October, 1999.

[12] T. Naito, T. Tsukada, K. Yamada, K. Kozuka and S. Yamamoto, "License Plate Recognition Method for Inclined Plates Outdoors", in: *Proceedings of the 1999 International Conference on Information Intelligence and Systems*, pp. 304-312, 1999.

[13] J.A.G. Nijhuis, M.H. ter Brugge, K.A. Helmholt, J.P.W. Pluim, L. Spaanenburg, R.S. Veenstra and M.A. Westenberg, "Car License Plate Recognition with Neural Networks and Fuzzy Logic", in: *Proceedings IEEE International Conference on Neural Networks*, Vol. 5, pp. 2232-2236, 1995.

[14] J.A.G. Nijhuis, M.H. ter Brugge, K. Kok and L. Spaanenburg, "Handling Unseen Data in a Neural License Plate Recognition System", in: *Proceedings EUFIT'97*, Vol. 1, pp. 380-383, September 8-11, 1997.

[15] R. Parisi, E.D. di Claudio, G. Lucarelli and G. Orlandi, "Car Plate Recognition by Neural Networks and Image Processing", in: *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems (ISCAS'98)*, Monterey, CA, Vol. 3, pp. 195-198, 31 May - 3 June, 1998.

[16] M. Preisenberger, F. Raupp, S. Schiessl and M. Hoffmann, "Integrierte Anwendung von Bildverarbeitung und Neuronalen Netzen zur Zeichenerkennung", in: *Prozessrechensysteme '91 (Proceedings Automatisierungs- und Leitsysteme in den neunziger Jahren, Berlin, Germany, 25-27 February, 1991)*, Springer Verlag, Berlin, Germany, pp. 350-363, 1991.

[17] L. Salgado, J.M. Menéndez, E. Rendón and N. García, "Automatic Car Plate Detection and Recognition Through Intelligent Vision Engineering", in: *Proceedings IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology*, Madrid, Spain, pp. 71-76, 5-7 October, 1999.

[18] K. Tanabe, E. Marubayashi, H. Kawashima, T. Nakanishi and A. Shio, "PC-Based Car License Plate Reading", in: *Proceedings of the SPIE (Image and Video Processing II, San Jose, CA, 7-9 February, 1994)*, Vol. 2182, pp. 220-231, 1994.

[19] A.E. Wiggins, "Helsinki Journey Time Monitoring System", in: *Proceedings IEE Seminar on CCTV and Road Surveillance*, London, UK, pp. 7/1-7/9, 12 May, 1999.

[20] N. Zimic, J. Ficzko, M. Mraz and J. Virant, "The Fuzzy Logic Approach to the Car Number Plate Locating Problem", in: *Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems (IIS'97)*, pp. 227-230, 1997.