PIXEL CLASSIFICATION FOR AUTOMATED DIABETIC FOOT DIAGNOSIS

C. Kloeze¹, A. Klein¹, S. Hazenberg², F. van der Heijden¹, J.G. van Baal², S.A. Bus²

¹University of Twente, Signals and Systems group, The Netherlands
²Ziekenhuisgroep Twente, location Almelo, The Netherlands

1 Introduction
Worldwide, more than 180 million people suffer from diabetes mellitus and this number is growing rapidly (World Health Organization). Approximately 50% of these patients will develop complications to their feet [2,7]. Diabetes mellitus is a disease in which the blood sugar becomes too high, damaging the blood vessels, nerves and other internal structures of the body. Due to this damage, the blood supply to the skin and nerves is reduced, causing neuropathy. This results in loss of protective sensation in the feet. Combined with poor blood supply and biomechanical changes, this results in a high risk for foot ulcers, which is a key problem in the diabetic foot.

Other important signs of diabetic foot disease are abundant callus formation, fissured skin, redness of the skin, blisters and increased local skin temperature. Some examples of a diabetic foot are shown in figure 1.

![Fig.1: Examples of diabetic feet](image)

These chronic wounds can become infected and ultimately lead to lower extremity amputation, which has a serious effect on the quality of life of the patient, and causes a large economic burden on society. Diabetes is the most common cause of nontraumatic lower extremity amputations in the industrialized world [1]. However, early detection and adequate treatment of the foot complications such as ulcers may prevent many of these amputations [1]. It is thus important that patients are checked regularly, preferably on a daily basis.

This was the motivation for a collaborate project (Vincent50) in which a photographic foot imaging device was developed. This device is developed to produce photographic images of the sole of the foot in the patient’s home which are automatically sent to a database server from where they are downloaded and assessed by clinical experts. The system allows scanning of the foot soles on a daily basis which may lead to early recognition of foot problems [4,5,6]. Currently, this device has been tested in the home of 25 patients.

The goal of the present study is to determine whether pixel classification is a useful intermediate step towards automatically assessing the images of the foot soles for signs of diabetic foot disease. If successful, this approach will further relief health care professionals in assessing the foot and enable the placement of more devices in the future.

2 Methods
In this study, the classification paradigm was used to design algorithms which can classify pixels for signs of diabetic foot disease. These algorithms were validated by comparing the results to annotations done by medical experts on diabetic feet. All algorithms were designed using Matlab r2007b and the additional toolbox PRTtools4.

Pixel classification is an application of statistical pattern classification, in which a class label is assigned to the individual pixels of an image of an object. The basic principle of pixel classification is shown in figure 2 [8,9].

The objects in this study were images of the soles of the feet of 32 diabetic patients with a variety of foot problems. These images show signs of foot disease of four possible classes: ‘ulcer’, ‘callus’, ‘fissure’ and ‘healthy’. Not enough information was available in the data to use the classes ‘redness’ and ‘blisters’ in the classification process.
2.1 Annotations

All images were annotated by a surgeon with several years of experience with treating diabetic feet to provide a reference. These annotations were translated from paper versions to digital annotations per pixel (the reference labels) by two experts. The interobserver reliability was calculated and Cohen’s kappa was found to be 0.7, which is considered good. The annotations of one expert were used as a reference to assess the designed algorithms and to construct the feature vectors.

2.2 Foot pixel classification

It is important in the classification process that only foot pixels are taken into account. The foot pixels were separated from the background by using pixel classification. Three classes were used: ‘foot’, ‘background’ and ‘beam’ (i.e. the mediolateral foot support used for positioning the foot in the device). Using a quadratic classifier, the unbiased error performing a 5-fold crossvalidation was found to be 2%. An example of the foot pixel classification can be seen in figure 3.

The features were chosen such that they result in different values for the four diabetic foot sign classes. Based on these differences the classes can be distinguished from each other by a classifier.

2.3 Features and classifiers

Three different types of features were designed to extract information from the foot pixels in the images: color (RGB), first order Gaussian derivative features, and second order Gaussian derivative features. The RGB features represent the red, green and blue values of the pixels. The first order Gaussian derivatives were calculated with different scales (sigma is 1, 2 and 3). The second order Gaussian derivative (the Laplacian) was calculated for four different scales: sigma is 2, 3, 4 and 20. In figure 4 the Laplacians with a scale of 4 and 20 of two feet are shown.

The features were used to make a labeled dataset, in which the reference labels were combined with the feature information. Six datasets were constructed with different combinations of features: RGB, 1st order, 2nd order, RGB and 1st order, RGB and 2nd order, and a combination of all three types of features. Per class 1000 pixels were randomly selected from all pixels available in that class, resulting in a dataset with a size of 4000 x number of features.

These labeled datasets were used to train four classifiers: a linear classifier (ldc), a quadratic classifier (qdc), a k-nearest neighbour classifier (knc) and a Parzen classifier (parzen). The trained classifiers were applied to the original images to assign an estimated class to each pixel in the image.

2.4 Performance diabetic feet classification

To assess the performances of the combinations of features and classifiers a 5- fold crossvalidation was
performed and Cohen’s kappa was calculated. With a 5-fold crossvalidation the unbiased error was calculated, resulting in a value between 0 (no error) and 1 (maximum error). By calculating Cohen’s kappa, the agreement between the assigned labels and reference labels was calculated, resulting in a value between 0 (no agreement) and 1 (perfect agreement). In medical applications a kappa between 0.4 and 0.7 is considered good [3], which we also consider good for this study.

3 Results & Discussion

The results of the performances of features and classifiers in terms of Cohen’s kappa and the 5-fold crossvalidation are shown in figure 5.

Some differences in performances can be seen between Cohen’s kappa and the error. This can be explained by the fact that Cohen’s kappa was calculated without application of cross-validation. That is, training and evaluation were performed with the same set (possibly leading to an optimistically biased performance). In contrast, the error rate was calculated with cross-validation included. Thus, overfitting is not likely to occur here.

The overall tendency that can be seen in figure 5 is that Cohen’s kappa and the error both improve when more features are combined. Both graphs show that the k-nearest neighbour and the Parzen classifiers are overall the best performing classifiers. The linear classifier is the least performing classifier in all combinations of features.

The RGB features are a solid basis for diabetic foot classification. The other individual features (1\text{st} and 2\text{nd} order Gaussian derivative features) perform less. However, in combination with the RGB features the performance does increase in comparison to the performance of the individual RGB features.

The optimal combination of features and classifier proved to be the combination of RGB, 1\text{st} and 2\text{nd} order Gaussian derivative features combined with the Parzen classifier. Cohen’s kappa was found to be 0.62 (which is a good agreement), and 81% of the pixels representing ulcer were correctly assigned. It is important for our application that as much as possible ulcer pixels are correctly assigned, due to the risk of infection, and eventually lower extremity amputation in these cases.

Two examples of the diabetic foot sign classification are shown in figure 6.

Both images show the contours of the diabetic foot signs present in the reference assessment. It can also be seen in these classified images that a relatively large amount of the healthy

---

IEEE-EMBS Benelux Chapter Symposium November 9-10, 2009
pixels in the reference assessment are assigned to one of the diabetic foot sign classes, or that ‘callus’ and ‘fissure’ pixels are sometimes mixed up in the classification process. Also the edges of the foot are assigned to being ulcer, which might be explained by the photometric settings of the light when the photographic images of the foot soles were made. This needs to be improved in further research.

The results of this study were influenced by some aspects. Firstly, the results are influenced by the translation of the annotations on paper to reference labels. Cohen’s kappa was calculated between two experts who performed the translation and was found to be 0.7. This is considered good, but it is also a fact that errors in some amount are made in this process. These errors influence the result of the classification because the reference labels form the basis of the classification process. Secondly, the computational complexity of the k-nearest neighbour and Parzen classifiers forced us to make relatively small datasets. Thirdly, the relatively small size of the available image database can also have had influence on the results. Because of the large variability present among the feet that were imaged it is important to have a large set of images to train the classifiers. More images of diabetic feet are necessary to get more reliable results in further research.

In this study, we applied only some rudimentary classifiers. Other classifiers like neural networks or support vector systems might also improve the results. Furthermore, we used the same priors and costs for all classes. In diabetic foot classification it is important that the different signs of diabetic foot disease are recognized, in particular the ulcer pixels. By, for example, raising the prior for finding an ulcer pixel in the process, we believe the result will improve.

4 Conclusion & Recommendations

A first step is taken in the automation of diabetic foot diagnosis from photographic images taken of the sole of the foot. The best agreement between automated recognition and expert diagnosis was achieved with a combination of RGB and derived features, proves that the RGB data is informative with respect to detection of ulcers. However, the automatic detection of pre-signs of ulcers and other anomalies needs more sophistication than pixel classification alone. Below, some recommendations for further research are given.

First of all, in this study, the usage of RGB data stems form the fact that the current foot imaging device is designed for manual processing of photographs. Other physical features, such as hyperspectral data, infrared and/or textural features are expected to be more informative (rich textural features can be scanned by using photometric stereo). Secondly, in this study we treated the pixels individually. We expect to be able to boost the performance by using the context between neighboring pixels. Thirdly, an individualized and normalized classification process might help with the large variability in foot soles between individuals. Additionally, when the classification is only based on foot sole images made in the history of an individual, we suspect that the classification will give more reliable results.

References