SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA
FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY

Ing. Jozef Goga

Dissertation Thesis Abstract

DIAGNOSTICS IN BIOMEDICINE BASED ON NEURAL NETWORKS

to obtain the Academic Title of

„philosophiae doctor“, abbreviated as „PhD.“

In the doctorate degree study programme:	Robotics and Cybernetics
In the field of study:	9.2.7. Cybernetics
Form of study:	full-time
Place and date:	Bratislava, 28.9.2022

Dissertation Thesis has been prepared at

Institute of Robotics and Cybernetics, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, Ilkovičova 3, 812 19 Bratislava, Slovak Republic

Submitter: **Ing. Jozef Goga**
Institute of Robotics and Cybernetics
Faculty of Electrical Engineering and Information Technology
Slovak University of Technology in Bratislava
Ilkovičova 3, 812 19 Bratislava, Slovak Republic

Supervisor: **prof. Ing. Jarmila Pavlovičová, PhD.**
Institute of Robotics and Cybernetics
Faculty of Electrical Engineering and Information Technology
Slovak University of Technology in Bratislava
Ilkovičova 3, 812 19 Bratislava, Slovak Republic

Readers: **prof. Ing. Igor Farkaš, Dr.**
Department of Applied Informatics
Faculty of Mathematics, Physics and Informatics
Comenius University
Mlynská dolina, 842 48 Bratislava, Slovak Republic

doc. Mgr. Radim Krupička, Ph.D.
Department of Biomedical Informatics
Faculty of Biomedical Engineering
Czech Technical University in Prague
nám. Sítná 3105, 272 01 Kladno 2, Czech Republic

Dissertation Thesis Abstract was sent:

Dissertation Thesis Defence will be held on:

at Faculty of Electrical Engineering and Information Technology
Slovak University of Technology in Bratislava
Ilkovičova 3, 812 19 Bratislava, Slovak Republic

prof. Dr. Ing. Miloš Oravec
Dean of FEI STU

SÚHRN

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Študijný program:	Robotika a Kybernetika
Autor:	Ing. Jozef Goga
Dizertačná práca:	Diagnostika v biomedicíne s využitím neurónových sietí
Vedúci záverečnej práce:	Prof. Ing. Jarmila Pavlovičová, PhD.
Miesto a rok predloženia práce:	Bratislava 2022

Hoci efektívne využívanie nových metód v oblasti biomedicíny prináša výrazné vylepšenia, reálne aplikácie sú stále výzvou, najmä z dôvodu zložitosti medicínskych dát. Vzhľadom na dôležitosť medicínskej diagnostiky sa v tejto oblasti zavádzajú inteligentné metódy, ktoré viedli v posledných rokoch k vytvoreniu novej vednej oblasti - umelej inteligencie v medicíne. Z dôvodu komplexnosti riešenej problematiky ako celku sa v tejto práci zameriavame na analýzu špecifických podoblastí medicínskej diagnostiky s cieľom zlepšiť existujúce metódy a odstrániť nedostatky aktuálne používaných modelov na báze umelých neurónových sietí. Z mnohých rôznych metód sa naša pozornosť upriamuje na biologicky inšpirované prístupy teórie viacerých rozlíšení a waveletovej analýzy, ktoré sú vhodné na extrakciu kontextových a štrukturálnych informácií z vizuálnych dát. Inkorporáciou týchto prístupov do najnovších modelov neurónových sietí vo forme fixných filtrov určených na extrakciu intrinzických príznakov sme analyzovali zmeny v architektúre siete schopnej efektívneho učenia aj pri limitovanom počte trénovacích vzoriek. Pri skríningu fundus obrazov sietnice pomocou hlbokých neurónových sietí sme zavedením stroja s podpornými vektormi vo forme pomocného klasifikátora zlepšili a urýchlili proces lokalizácie tvrdých exudátov ako príznakov diabetickej retinopatie. Za účelom širšieho uplatnenia biometrických systémov v biomedicíne sme na základe rozsiahlej analýzy odpovedali na šesť vedeckých otázok obsahujúcich odporúčania pre pokračovanie výskumu v oblasti rozpoznávania ľudských dúhoviek s cieľom zlepšiť kvalitu a dostupnosť verejných databáz.

Kľúčové slová: konvolučné neurónové siete, diabetická retinopatia, rozpoznávanie dúhovky, hlboké učenie

ABSTRACT

SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA
FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY

Study Programme:	Robotics and Cybernetics
Author:	Ing. Jozef Goga
Dissertation:	Diagnostics in Biomedicine Based on Neural Networks
Supervisor:	Prof. Ing. Jarmila Pavlovičová, PhD.
Place and year of submission:	Bratislava 2022

Although the effective use of novel methods in biomedicine brings significant improvements, real-world applications are still a challenge, especially as regards the complexity of medical data. Given the importance of medical diagnostics, intelligent methods are being introduced in this area, which in recent years have led to creation of a new scientific field - artificial intelligence in medicine. Due to the complexity of the entire field, in this work we focus on the analysis of specific sub-areas of medical diagnostics in order to improve existing methods and eliminate the shortcomings of currently used models based on artificial neural networks. Of the many different methods, our attention is focused on biologically inspired approaches such as scale-space theory and wavelet analysis, which are suitable for extracting contextual and structural information from visual data. By incorporating these approaches into the latest models of neural networks in the form of fixed filters designed for the extraction of intrinsic features, we created and analysed changes in the network architecture capable of effective learning even with a limited number of training samples. In the screening of retinal fundus images using deep neural networks, we have improved and accelerated the process of localisation of hard exudates as features of diabetic retinopathy by introducing a support vector machine in the form of an auxiliary classifier. For the purpose of a wider application of biometric systems in biomedicine, based on an extensive analysis, we answered six scientific questions containing recommendations for the continuation of research in the field of human iris recognition with the aim of improving the quality and availability of public databases.

Keywords: convolutional neural networks, diabetic retinopathy, iris recognition, deep learning

Contents

CHAPTER 1

Introduction

Several successful healthcare applications in recent years have shown the undisputed benefits of using artificial intelligence as an appropriate tool to solve complex problems. In the field of healthcare, especially in biomedicine, in addition to a specific non-trivial problem, expertise is also required. Despite the availability of a considerable amount of data, conventional methods often fail mainly due to the difficulty and nonlinear nature of the problems addressed [1]. Therefore, the role of the so-called expert is increasingly being taken up by complex computational models such as artificial neural networks. However, these are not applied as the only decision-making system in the diagnosis process. In addition to legal aspects taking into account strict regulations requiring clinical trials prior to the introduction of given systems into medical practise, the primary reason is always patient safety. Due to the complexity of these computational models, it is not always clear which parameters influence the current output hypothesis. In the event of an error, life-threatening situations could arise, and therefore it is necessary to assume responsibility for them. However, in the case of artificial neural networks, this issue is so complex that current legislation has not yet answered these and many other emerging questions [2]. In spite of these shortcomings, artificial neural networks are increasingly more popular in medicine, largely due to their high accuracy, which exceeds standard computational models. However, even despite the people's awareness of their undeniable advantages, they often face cultural resistance. Mistrust in new technology can only be overcome by practical applications and achieving better results compared to the current state regardless of the area concerned. And that is what we are witnessing in the field of biomedicine today. In this respect, it is important to analyse the improvement that can bring the application of these computational models to medical practise. What consequences and shortcomings are associated with their deployment and, ultimately, why is it important for research in this area to proceed.

CHAPTER 2

Improving Neural Network Models via Visual Priors

In the field of biomedicine, the most commonly used convolutional neural networks are inspired by the well-known research of Hubel and Wiesel [3], who investigated the primary visual cortex (V1) of cats by analysing the neural response to various visual stimuli. These have undergone significant progress in recent years, aimed at eliminating shortcomings and deepening architectures in order to increase their overall accuracy on standardised benchmarks. In addition to usual experiments, visualisation methods have been developed to better understand the processing of information within their architecture [4, 5]. These revealed a parallel with biological findings in neuroscience, which points to a gradual distinction between simple shapes and textures in the first layers of the neural pathway and complex shapes and objects in the deeper layers [6]. An example is the common generation of visually similar convolution kernels resembling Gabor filters in the first layer, suggesting that convolutional neural networks mimic the biological processing of visual information at the fundamental level. Despite these discoveries, major developments in deep neural networks focus on improving their accuracy, and achieving a correspondence with neuroscience is secondary. In recent years, however, some scientists have expressed views on the simplification of deep neural networks in order to remove the black-box principle, which would lead to a deeper understanding and thus improved models. One emerging direction of research is to introduce visual priors into architectures in order to understand their impact not only on the accuracy achieved, but also on the creation of features within a given application domain.

We can decompose regular convolution into two simpler operations: extraction of features in the spatial plane using dephtwise convolution and subsequent channel recombination using pointwise convolution effectively creating a depthwise separable convolutional layer [7, 8]. Depthwise convolution is a convolution in the spatial domain that is applied independently over every channel of an input. The calculation of the k-th feature map $\tilde{\mathbf{h}}_k^y \in \mathbb{R}^{1\times N_K D_x\times(W_x-W_K+1)\times(H_x-H_K+1)}, k \in \{0,1,2,\ldots,N_K\}$ by 2D depthwise convolution is described by Equation 1. The output is a three-dimensional tensor, where

the index $o \in \{0, 1, 2, \ldots, N_K D_x\}$ represents a specific channel in the depth plane and $p \in \{0, 1, 2, \ldots, H_x - H_K\}$ and $q \in \{0, 1, 2, \ldots, W_x - W_K\}$ are their vertical and horizontal indices.

$$\begin{aligned} \tilde{\mathbf{h}}_k^y(o,p,q) &= 2DDepthwiseConv\left(\mathbf{h}^x, \mathbf{\Theta}_k\right)(o,p,q) \\ &= \sum_{u=0}^{H_K-1} \sum_{v=0}^{W_K-1} \mathbf{h}^x(i, p+u, q+v) \cdot \mathbf{\Theta}_k(H_K-u-1, W_K-v-1) \end{aligned} \tag{1}$$

Subsequent pointwise convolution can be seen as regular convolution with 1×1 kernel size, recombining the channels along depth and projecting them onto a new channel space, creating an output j-th feature map $\mathbf{h}_j^y, j \in \{0, 1, 2, \ldots, D_Y\}$ (Eqn. 2).

$$\mathbf{h}_j^y(p,q) = PointwiseConv\left(\tilde{\mathbf{h}}_k^y, \boldsymbol{\alpha}_j\right)(p,q) = \sum_{o=0}^{N_K D_x - 1} \tilde{\mathbf{h}}_k^y(o,p,q) \cdot \boldsymbol{\alpha}_j(o) \tag{2}$$

The original set of parameters $\mathbf{K}$ is hereby replaced by $\mathbf{\Theta} \in \mathbb{R}^{N_K \times 1 \times W_K \times H_K}$ as a set of effective filters and $\boldsymbol{\alpha} \in \mathbb{R}^{D_Y \times N_K D_x \times 1 \times 1}$ for the subsequent recombination of the created feature maps. This is the main idea of the Receptive Fields Neural Network (RFNN), which replaces the parameters of standard convolutional kernels $\mathbf{\Theta}$ with the N_K fixed base filters $\boldsymbol{\theta}$. This makes it possible to introduce information into the network a priori in the form of fixed spatial 2D convolutional filters for the extraction of intrinsic features, which are then linearly combined creating effective feature maps $\mathbf{h}^y$ at the output of the RFNN layer (Eqn. 3).

$$\mathbf{h}^y = PointwiseConv\left(2DDepthwiseConv\left(\mathbf{h}^x, \boldsymbol{\theta}\right), \boldsymbol{\alpha}\right) \tag{3}$$

By introducing fixed filters from a certain domain, the convolutional kernels can be treated as functions, while the network parameters are also reduced, as only $\boldsymbol{\alpha}$ parameters are learnt during the training process.

2.1 Methods

As a baseline for comparative evaluation of RFNN networks, we selected the methodology and results of Jacobsen et al. [9] described below. We have reimplemented their proposed RFNN network to reproduce the original published results and ensure correct implementation using the Pytorch [10] framework. We refer to this implementation as $RFNN_{ref}$. Since the original article contains several experiments and architectural modifications, we

decided to concentrate on the results, where the MNIST [11] dataset was used (Figure 5 in the original article [9]). The baseline architecture is a convolutional neural network with three layers of receptive fields, which is used to classify the known problem of handwritten digits into ten classes.

Fixed convolution kernels from Gaussian basis were formed in the same way for all layers with one difference in size. In the first layer, the convolutional kernels have a size of 11×11 pixels and $\sigma = 1.5$, while in the second and third layers the kernels are identical with a size of 7×7 pixels with $\sigma = 1$. In each receptive fields layer, the filtered outputs after depthwise convolution are recombined into 64 output feature maps. Receptive fields neural networks excel when not enough data is available to use common deep learning models. We used a limited number of training samples in the range from 300 to 60,000 samples, i.e. the whole training set. For each experiment, the entire MNIST test set, consisting of 10,000 samples, was used to assess the accuracy of classification. In addition to normalising data to the range $< 0, 1 >$, no other form of sample preprocessing was used. The network was trained by a standard backpropagation algorithm using an average cross-entropy loss with a mini-batch size of 25 images. An Adadelta optimizer with decay rate $\rho = 0.95$, stability constant $\epsilon = 1e - 6$ and linear learning rate decay was applied. The initial value of the learning rate was set to $lr_i = 5$ and decreased to the final value of $lr_f = 0.05$ during the training, while a fixed number of epochs for individual experiments were taken from the original authors [9]. The resulting classification accuracy of the reference architecture $RFNN_{ref}$ representing the average of three experiments is depicted in Fig. 1. The original and our result lines approximately match, but there is apparent difference in the last point for the least training samples. The variability of these results even if the original methodology is followed may be due to various factors such as random number generator (RNG) initial state, hardware parameters, software versions of used frameworks, and others. These data are missing in most published studies [12], which makes the verification of the original results more difficult. Although we managed to achieve better results than the original authors, we noticed the shortcomings of such an evaluation, which is common in the field of neural networks.

The results may be affected by the setting of hyperparameters and training methodology, randomness of selection and order of samples, as well as stochastic regularization (e.g. dropout) during training. These effects are more pronounced with a smaller training set, which in our case was manifested by an increase in variance. Due to the difficulty of reproducing the original results, we tried to ensure reproducibility of our approach by setting the methodology as follows.

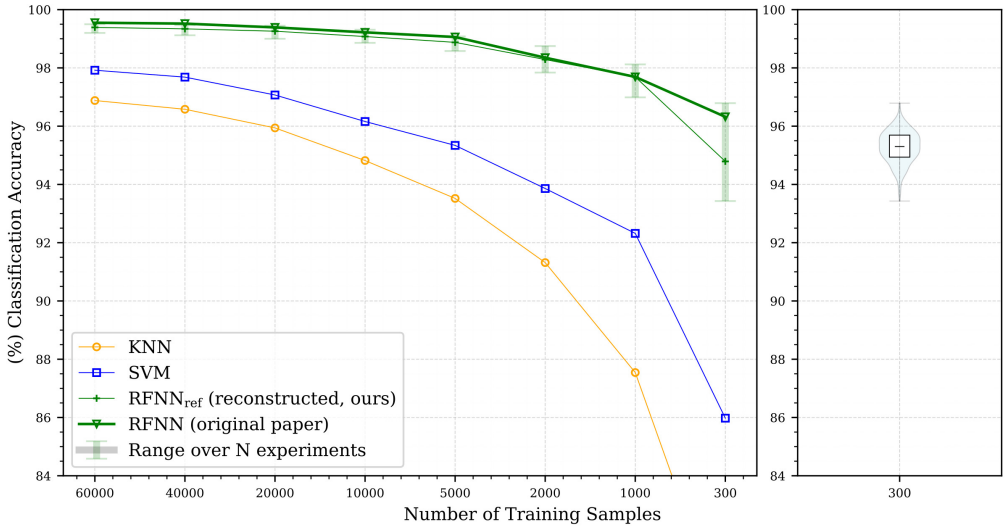

Figure 1: *Reproduction of the original results published on the MNIST database. The average of the three experiments is shown in green for all samples. Accuracies of the standard KNN and SVM machine learning classifiers are given as separate curves. The graph also shows the range of results obtained by repeating the experiment $N = 300$ times for the reference architecture $RFNN_{ref}$. For the smallest number of training samples, the estimated distribution of all repeated experimental results shown as a violin plot is also shown separately (best viewed in colour).*

- All used frameworks were set to deterministic mode
- All parameters were initialised based on one chosen experimental seed
- To ensure variability, the selected experimental seed was used to generate a seed vector for all runs of the experiment

To evaluate whether one neural network on a given dataset performs better than another one, we originally chose the Null Hypothesis Significance Testing (NHST) methodology, such as the Friedman and Nemenyi post-hoc tests [13, 14]. Based on recommendations [15, 16] we abandoned it and chose a Bayesian-based comparison in the form of the Correlated Bayesian t-test [17]. We used an implementation of this test from the Baycomp library [18]. The result of this test are three probabilities: the probability that the models are practically equivalent and the probabilities that model C1 is better than model C2 and vice versa. For all tests, we chose standard 1% as the Region of Practical Equivalence (ROPE), where with a smaller difference in the metrics of the compared models C1 and C2 as defined 1%, we consider the models as equivalent. Based on previous adjustments and evaluations, we decided to change the methodology of the experiments. Subsequently, in

each experiment, all frameworks using randomness are set to the deterministic mode. We used stratified 10-fold cross-validation repeated 10 times with a different initial seedings (a total of 100 simulations). The same master seed was used for each experiment, so not only the selection but also the order of the samples was maintained for all models on the same data set. The statistical evaluation of the model in the form of a correlated Bayesian t-test was performed on the validation part of the training set within the cross-validation. However, a numerical evaluation was performed on the held-out test set containing the same 10,000 samples to obtain an objective view of the real accuracy of the model. A possible alternative is to combine the train and the test set into one dataset. We chose not to use this procedure in order to compare the results obtained on a dedicated test set with previous research in this area.

2.2 Results

For a valid comparison of the impact of changes in the RFNN architecture, we simplified the reference network as much as possible. Due to the possible mutual influence of parameters in deep architectures, we omitted the normalization and regularization layers and simplified the neural network to a shallow architecture with one computational RFConv layer. Another simplification was the removal of computational layers (local response normalization, dropout), for which we could not ensure repeatability when training on a GPU. We kept the base and number of filters according to the original architecture. For training, we used the same settings as in the previous experiments, except that we changed the optimizer to AdamW using a fixed learning rate with a value of 1.10^{-3}. For reference, we evaluated the simplified single layer architecture $RFNN_{L1}$ against the original three-layered architecture $RFNN_{ref}$ in the form of 10 times repeated 10-fold cross-validation using the correlated Bayesian t-test for random selection of 300 samples. Compared to the original architecture, the results deteriorated significantly, which was expected. For a ROPE of 1% the probability that $RFNN_{ref}$ is better than $RFNN_{L1}$ was 99.03%, the probability that both models are practically equivalent was 0.95% and the probability that $RFNN_{L1}$ is better than $RFNN_{ref}$ was 0.02%.

Before making changes to the simplified architecture $RFNN_{L1}$, we analysed the impact of sample selection on the accuracy achieved by the created model. The setting and methodology of the training were the same as in the previous section. We chose $RFNN_{L1}$ as the reference architecture and conditioned the selection of samples by changing the deterministic seed that affected the result. We trained again on a random selection of 300 samples and re-evaluated classification accuracies using the same methodology.

Normalization in neural network architectures can help both to improve convergence and generalisation [19]. Different types of preprocessing and normalization are introduced to eliminate sharp differences between network parameters or extracted feature maps [20]. This effect can be even more pronounced when using fixed filters. For this reason, we evaluated the impact of energy normalization on the accuracy achieved. By energy normalization we mean normalization to unit energy, where the sum of all squared k-th filter elements $\boldsymbol{\theta}_k \in \mathbb{R}^{W_K \times H_K}$ is equal to one (Eqn. 4).

$$\boldsymbol{\theta}_k^{norm}(m,n) = \frac{\boldsymbol{\theta}_k(m,n)}{\sqrt{\sum_{u=0}^{H_K-1}\sum_{v=0}^{W_K-1}\boldsymbol{\theta}_k^2(u,v)}} \tag{4}$$

where $m \in \{0,1,2,\ldots,H_K\}$ and $n \in \{0,1,2,\ldots,W_K\}$ are their corresponding vertical and horizontal indices.

The results of both experiments show that base energy normalization helped improve classification accuracy. When using the original Gaussian derivative base, we found that $RFNN_{L1}(Gaussian)$ was better than $RFNN_{L1}(Gaussian\ normalised)$ with a probability of 6.75%, the probability that both models are practically equivalent was 55.06% and the probability that $RFNN_{L1}(Gaussian\ normalized)$ is better than $RFNN_{L1}(Gaussian)$ was 38.19%. In case of using a random base the probability that $RFNN_{L1}(Random)$ is better than $RFNN_{L1}(Random\ normalized)$ was 21.15%, the probability that both models are practically equivalent was 49.21%. And lastly, as we expected, the probability that $RFNN_{L1}(Random\ normalized)$ is better than $RFNN_{L1}(Random)$ was 29.64%.

In the original article [9], a specific feature of the RFNN is the Gaussian compact base, which is used for feature extraction. The aim of this experiment was to verify whether we can replace the given base and what effect this change will have on the achieved classification accuracy. We again chose $RFNN_{L1}$ as the reference architecture, while the parameters and evaluation methodology remained unchanged. In each experiment, we kept all parameters constant except for the base used with a size of 11×11 pixels. Since the reference Gaussian base contains 10 filters, we selected these filters along the primary diagonal based on a triangular selection according to Ulicny et al. [21]. We compared the reference Gaussian derivative base (Gaussian), the orthonormal Discrete Cosine base (DCTII), the orthonormal Discrete Hartley base (DHT), the random base with normal distribution (RND), and the random orthonormal base (ORTHN_RND).

We evaluated neural models using a Correlated Bayesian t-test with respect to the simplified reference architecture $RFNN_{L1}$ with a Gaussian base for each model change

separately. The results showed that for all tested bases the hypothesis that both compared models are equivalent has the highest probability with the one exception of random base, which worsened the results and had the greatest impact on achieved classification accuracy. The probability that the $RFNN_{L1}$ model of a network with a Gaussian base is better than a model with a random base was about 47,1 %, far exceeding other results. Compared to the other bases tested, we obtained the best results using a discrete Hartley and discrete cosine base. In both cases, the result of the statistical test is that the models are practically equivalent with a 51.06% and 44.77% probability, respectively.

Table 1: Results of the test classification accuracies obtained using a simplified $RFNN_{L1}$ architecture with various fixed bases of the same size trained using only 300 samples from the MNIST database.

Base	Mean Accuracy [%]	Min Accuracy [%]	Max Accuracy [%]	Variance [%^{2}]
Gaussian	89.67	86.30	91.75	1.0131
Discrete Cosine (DCTII)	**90.21**	88.45	**91.86**	0.7268
Discrete Hartley (DHT)	90.20	**88.47**	91.76	0.5958
Random (RND)	88.48	86.15	90.09	0.9813
Orthonormal random (ORTHN_RND)	89.32	87.36	90.65	**0.5528**

2.3 Discussion

In this work we analysed the receptive field neural network (RFNN), as a promising simplified model, which uses a Gaussian derivative basis inspired by the scale-space theory. When reconstructing the original results, we identified a problem with a large variation of the achieved test classification accuracy, especially with a small amount of training data. We modified the training methodology to ensure repeatability and evaluated proposed step-by-step changes in the form of a Correlated Bayesian t-test using 10 times repeated 10-fold cross-validation. Using this methodology, we created a single-layer CNN network inspired by the RFNN architecture, whose input of computational layers is a set of 2D fixed filters of any size and number. Subsequently, we experimentally verified that bases other than Gaussian can be used to generate fixed filters within the created architecture. We found that a change in base may have less of an effect on the results obtained than retraining the network when using another seed by failing to ensure repeatability of results. We also verified the positive impact of energy normalization of used filters, which significantly improves the results obtained, even when using a random base.

CHAPTER 3

Localisation of Hard Exudates Based on Deep Learning

Given the extent and severity of diabetic retinopathy, early detection and treatment are one of the key goals of most healthcare facilities around the world [22]. Therefore, automating the screening of patients at risk of developing this disease represents not only an opportunity for early diagnosis, but also for reducing the direct and indirect costs of treating and addressing the complications associated with endangered patients at a later stage of the disease [23]. To compare the accuracy of existing methods, more than 40 databases have been created containing fundus and optical coherence tomography (OCT) images of the retina, the vast majority of which are publicly available. However, images in these databases are often captured by multiple sensors, markings are incomplete or incorrect, and of the relatively large number of databases available, only a small portion contains lesion-by-pixel annotations [24]. As a result, most research focusses on grading diabetic retinopathy without explicit detection of individual lesions on fundus images [25]. In this work, we therefore focus on the localisation of hard exudates in retinal fundus images with the aim of fast and robust visual pre-scanning, which can help the early diagnosis and treatment of diabetic retinopathy. When designing the proposed method, we focused on improving the localisation of hard exudates by preprocessing and simplifying the problem solved by deep neural network in order to increase the overall accuracy. Although we could analyse entire fundus images, due to computational complexity and preservation of details, we decided to work with image patches. These were created from the original images by cutting on patches of equal width and height with a shift of half the size of the patch in all directions. If the image size is not divisible by the patch size, the original image is padded with black borders to prevent the loss of information that would occur by omitting edge patches. On the basis of the ground truth markings at the pixel level, bounding boxes with a minimum area are subsequently created so that the individual lesions are clearly distinguished. If the patch does not contain exudates, it is discarded and will not be used in the training process. The remaining labelled lesion patches are then preprocessed using offline augmentation to create the final training

dataset. Since patches without exudates are not used when training the deep object detector, it is necessary to dynamically filter them out when analysing the image. In the proposed solution, the SVM classifier performs this task. The input image is first cut into image patches with a standard size of 224×224 pixels - these do not enter the SVM classifier directly, but are fed to the input of a deep neural network with ResNet-50 [26] architecture. Using the transfer learning approach, we applied the pre-trained ResNet-50 model on the ImageNet [27] database as a feature extractor, which also reduces dimensionality for the SVM classifier. As input to the SVM, we used a $1 \times 1 \times 2048$ feature vector from the last average pooling layer obtained by traversing input 224×224 pixel image patches through the pre-trained ResNet-50 model. In principle, a method other than SVM can also be used to solve this binary classification problem; in our implementation, we used linear SVM due to the high accuracy and speed of execution, which is also an important factor. In theory, by pre-scanning and filtering image patches without exudates it allows the deep object detector to focus on finding and detecting hard exudates in more detail, which manifests itself in increased accuracy. As detector, we used a faster region-based convolutional neural network (Faster-RCNN) [28] again with a ResNet-50 backbone architecture, which was trained only on positive samples. Faster-RCNN, replaces the previously used selective search method by incorporating a region proposal network (RPN), a fully convolutional network for predicting object boundaries and concluding their probabilities at each position, directly in the unified architecture using the attention mechanism. The RPN uses anchor boxes, pre-defined rectangular boundaries centred at the sliding window of multiple scales and aspect ratios, which depend on the searched objects in the image. At each location of the sliding window, multiple region proposals are simultaneously predicted leading to compelling acceleration and improved localisation.

3.1 Methods

To perform the experiments, we searched for a database with sufficient diversity and accurate labels at the pixel level. We chose the e-optha-EX database [29], which contains images with marked exudates and images without lesions of various resolutions. We deal with mixed image resolutions by dividing the original fundus images into patches of 224×224 pixels. This size was not chosen at random, but matches the input layer size of the ResNet-50 network, which is used for feature extraction, as well as the backbone architecture of the Faster-RCNN object detector. Thus, we did not need to change the resolution of the input images by resizing. Instead, the original fundus images were com-

plemented with black pixel borders in the horizontal and vertical directions. The entire e-optha-EX database was used in the experiments, which we first divided into images containing (47 images in total) and not containing exudates (35 images in total) according to the provided labels. Each group was then randomly divided into five approximately equal folds for 5-fold cross-validation. The division of the 47 images containing exudates was divided into three folds containing 9 images and two folds containing 10 images each. Dividing the image samples without exudates was easier, where each fold contained exactly 7 images. We maintained this random division by providing the same seed for the pseudo-random number generator in all subsequent experiments. We also did this to avoid methodological errors so that image samples from one patient would not be in the train and test set simultaneously and also to train the Faster-RCNN detector and the SVM classifier on the same data. By dividing the original images, we created a total of 307 image patches containing exudates. Using a shift of 112 pixels (half of the patch size) in the horizontal and vertical directions, 882 samples were generated. To expand the dataset even further, we created new image samples together with bounding boxes by mirroring the original and shifted image patches along the x axis, y axis, and both axes sequentially, creating additional 3567 samples. In this way, we created a total of 4756 colour 224×224 pixel image patches, which formed the final dataset. Most of these patches contained approximately 3 or more exudates. From the original images, we generated twice the number of image patches that did not contain exudates in the same way - these were used only when training the SVM classifier. When generating the data, we took care to maintain consistency according to the classification of individual images into individual folds within 5-fold cross-validation. Image patches without exudates were used only in the training of the SVM classifier, which is used in the task of fast image prescanning. The SVM classifier uses extracted features from the ResNet-50 network with the use of transfer learning, as its weights are constant and pre-trained on the ImageNet database. The 2048-dimensional feature vector was extracted from the last average pooling layer of the ResNet-50 network and used as input for the linear SVM classifier. In order for SVM to be able to effectively filter out negative samples, we used the above-mentioned division for the training, while we added 1 pure black image patch to each fold in order to improve recognition in the marginal parts of the fundus images. The SVM classifier training was always performed on $N - 1$ folds and the evaluation on the remaining fold, which was not used during training. We maintained the same evaluation methodology when training the Faster-RCNN detector, with the only difference that we used only patches containing exudates in order to directly localise them.

To select the number of clusters k and thus the number of anchor boxes, we used the methodology proposed by Redmon et al. [30]. We performed a series of k-means experiments, where we increased the number of clusters k and evaluated the mean IoU for each centroid. Based on the results, we chose $k = 7$ with the mean $IoU = 0.7374$ as a suitable compromise between computational complexity and the representation of various bounding box shapes in the training dataset. In the training process, we minimised the multi-part loss function proposed by the authors of the YOLO detector [31] consisting of the sum of classification loss, confidence loss and localisation error expressed as mean squared error (MSE). As an optimizer, we used a stochastic gradient descent with momentum (SGDM) with a learning rate 3.10^{-4} and a default momentum decay factor of 0.9. Due to memory requirements, we limited the size of the minibatch to 2 image patches and trained object detector for a fixed number of 10 epochs in all experiments. We determined this value experimentally by training a Faster-RCNN detector using an early stopping on the first fold of a cross-validation split with a changed seed for a pseudo-random generator. We stopped training if the value of the loss function on the validation set did not improve over 5 epochs. To obtain the highest possible sensitivity, we set the cutoff value for the negative overlap to 0.1, which means that prediction bounding boxes with a smaller IoU value are not considered correctly localised. To verify robustness, we used only the original patches during training or we also added online contrast augmentation (CA) of image patches. This was realised in the HSV colour space by a random change in the range of $0 - 20\%$ of the maximum hue and brightness value. We maintained the shade of the black background by thresholding every pixel with value below 10 concurrently in all channels, and after the random augmentation, the image was transformed back into the original RGB colour space. By default, the evaluation of localisation accuracy can be performed by counting individual correctly and incorrectly classified pixels, also referred to as pixel-level evaluation. This evaluation leads to various contradictory situations where some pixels are marked as true positives (intersection of ground truth and prediction), and other pixels of the same exudate are then marked as false positives or false negatives according to the specific situation of inaccurate overlap. Therefore, the same authors [32, 33] decided to change the evaluation methodology and considered the whole connected component to be marked as true positives only if it touches the ground truth. Zhang et al. [34] noticed that good results can also be achieved by creating a large mask that covers the entire set of ground truths. To prevent this, they proposed a hybrid method that requires minimal overlap σ between prediction A and ground truth B, which we used in this work and referred to as evaluation at the exudate level. This evaluation

consisted of setting the cutoff overlap at 20% ($\sigma = 0.2$), which distinguished between correct and incorrect exudate localisation. In this case, all exudate pixels are marked as true positives if they overlap with the ground truth area greater than 20% . For overlaps smaller than the threshold σ, only pixels in the intersection area ($A \cap B$) are marked as true positives (TP). Pixels in the remaining area of the prediction bounding box ($A \cap \bar{B}$), which did not overlap with the ground truth, were subsequently marked as false positives (FP). Vice versa, pixels in the remaining area of the ground truth ($B \cap \bar{A}$), which did not overlap with the prediction boundary box, were marked as false negatives (FN). All other pixels were considered true negatives (TN). If the predicted and ground-truth bounding boxes did not overlap, the corresponding pixel designations were unambiguous, and the standard metrics were subsequently evaluated from these values. In a clinical setting, it is important to evaluate the probability of findings in the entire image [35]. For this reason, we also evaluated the accuracy of localisation at the image level, using the prediction probability of individual findings at the output of the Faster-RCNN network from the image patches. If the maximum probability of separate findings was higher or equal to 0.9, we considered the entire image to be a positive sample. Otherwise, if the certainty of the findings were below this value, we marked the entire image as a negative sample. To independently evaluate the accuracy of the proposed method, it was also applied to the DiaretDB1 [36] and Messidor [37] datasets, which were not used during the training phase, and subsequently compared the results achieved with the previously published results of relevant authors.

3.2 Results

Evaluation at both exudate and image level was performed in the form of 5-fold cross-validation by comparison with the baseline Faster-RCNN with ResNet-50 architecture without SVM pre-scanning, as well as with and without online contrast augmentation. The results obtained on the validation set are shown in Table 2.

The results in Table 2 show that the proposed method improves overall localisation accuracy, which is also reflected in improvements in the F1 score, positive predictive value and specificity at both levels of evaluation. It is also clear that SVM pre-scanning results in a significant 23.4% decrease in FPR at the exudate and approximately 23.1% reduction at the image level compared to baseline. With the introduction of online augmentation during training, these differences are even more prominent, reaching up to about 30.8 % at the image and even 44.7% at the exudate level. However, not all metrics point to improvement. In the evaluation, we observed a decrease in sensitivity of 2 to 7% due to

Table 2: *Comparison of the proposed method with baseline Faster RCNN detector on e-optha-EX database with and without contrast augmentation at the exudate and image levels. The results displayed represent the mean value obtained on the validation set ± standard deviation.*

	Method	Sensitivity	Specificity	Accuracy	F1-score	PPV	FPR	FNR
Exudate level	Faster RCNN	0.8987 ± 0.0752	0.9953 ± 0.0019	0.9947 ± 0.0017	0.6781 ± 0.0378	0.5984 ± 0.0593	0.0047 ± 0.0019	0.1013 ± 0.0752
	Faster RCNN + SVM	0.8805 ± 0.0672	0.9964 ± 0.0012	0.9957 ± 0.0012	0.69262 ± 0.0462	0.6236 ± 0.06208	0.0036 ± 0.0012	0.1195 ± 0.0672
	Faster RCNN + CA	0.8575 ± 0.0492	0.9963 ± 0.0016	0.9956 ± 0.0014	0.6577 ± 0.0103	0.6432 ± 0.1023	0.0037 ± 0.0016	0.1425 ± 0.0492
	Faster RCNN + CA + SVM	0.8344 ± 0.0442	0.9974 ± 0.0009	0.9966 ± 0.0009	0.7069 ± 0.0753	0.6626 ± 0.0982	0.0026 ± 0.0009	0.1656 ± 0.0442
Image level	Faster RCNN	1	0.6286 ± 0.0782	0.8412 ± 0.0346	0.8786 ± 0.0244	0.7841 ± 0.0392	0.3714 ± 0.07825	0
	Faster RCNN + SVM	1	0.7143 ± 0.1010	0.8779 ± 0.0418	0.9044 ± 0.0299	0.8266 ± 0.0508	0.2857 ± 0.1010	0
	Faster RCNN + CA	0.9578 ± 0.0579	0.4857 ± 0.2595	0.7566 ± 0.1129	0.8222 ± 0.0713	0.7254 ± 0.1000	0.5143 ± 0.2595	0.0422 ± 0.0579
	Faster RCNN + CA + SVM	0.9578 ± 0.0579	0.7429 ± 0.1863	0.8669 ± 0.0768	0.8940 ± 0.0550	0.8429 ± 0.0839	0.2571 ± 0.1863	0.0422 ± 0.0579

a notable increase in the false negative rate. The contribution of the SVM classifier in the prescanning task was also analysed by a separate evaluation. We achieved an average 84.7% classification accuracy on the validation set with a maximum of 92.2%. These values are satisfactory, but they still leave room for improvement. The SVM pre-scanning also brought a significant acceleration of the image evaluation process, but we did not directly evaluate the time saved because it was not the subject of interest. The proposed method generally achieved a more significant improvement at the image level compared to baseline. An example is the 23.1% improvement in the F1 score, which has increased to approximately 50% with the introduction of contrast augmentation. At the image level, we also evaluated the Area Under The Curve (AUC), which reached an ideal value of 1 on the e-optha-EX database and high values of approximately 0.97 and 0.88 on the DiaretDB1 and Messidor datasets, respectively. These results were obtained using the best performing model of the Faster RCNN network from the 5-fold cross-validation with SVM pre-scanning and contrast augmentation of the dataset during the training (Faster RCNN + SVM + CA in Table 2). The DiaretDB1 and Messidor databases were used for a hold-out evaluation, as they were not used during the training phase. The comparison at the image and exudate level with the respective authors show that we have managed to outperform the previously published results at the exudate level and the accuracy at the image level is comparable to other methods.

3.3 Discussion

In this work, we presented a novel method for localisation of hard exudates as symptoms of diabetic retinopathy on fundus images using a Faster-RCNN detector with an auxiliary SVM classifier designed for fast prescanning. Using this approach, it is possible to filter out image samples without findings and thus enable a deeper analysis of potentially positive samples with lesions using a deep neural network. Classification into positive and negative samples in the form of fast SVM pre-scanning was performed on the basis of extracted feature vector from ResNet-50 network pre-trained on the ImageNet database. The identification and localisation of hard exudates was based on a Faster-RCNN detector with the same architecture and was trained only on positive samples containing lesions. In the experiments, we tested four variants of the proposed method by comparison with the baseline architecture without the use of SVM pre-scanning and with and without contrast augmentation of the dataset. We experimentally verified the accuracy of the proposed method using five-fold cross-validation at two levels of evaluation: exudate and image level. Subsequently, we compared the achieved results with the authors of relevant articles dealing with the issue of localisation of hard exudates in fundus images. When evaluating the results at the exudate level, we recorded a 29.7% decrease in the false positive rate compared to the baseline at the cost of an increase of approximately 16.2% in the false negative rate. At the image level, we obtained even better results, namely, a 50% decrease in false negative rate without an obvious increase in false negatives. Although this method reduces sensitivity due to an increase in false negatives, these results may be acceptable due to a significant improvement in image-level detection, which is more emphasised in clinical practise. In terms of localisation accuracy, our approach outperformed or achieved comparable results with other methods, even in holdout evaluations on the DiaretDB1 and Messidor databases, which were not used during the training. Our analysis shows that prescanning images using SVM prior to deep learning detector can help improve the overall localisation accuracy achieved, as well as speed up the process of evaluating individual images, as indicated by the preliminary results.

CHAPTER 4

Iris Imaging in Healthcare Setting

Research in the field of iris recognition is constantly growing and has been further accelerated by technological advances in mobile devices in recent years. Along with the development of artificial intelligence, there is a need for robust recognition in the changing real-world conditions, which has led to the creation of new challenging databases to cover a wide range of noise factors and specific subareas of iris recognition. With the growing interest of the scientific community, the number of scientific publications, methods, and datasets has naturally increased as well, which means that more effort is needed to gain sufficient insight while researching. As this phase is often time-consuming, many researchers have decided to create their own private database instead of using publicly known databases or searching for available ones. There are several reasons. Due to the sensitive nature of the data, most databases are available for download only if a licence agreement, often requiring the signature of the legal representative of the research institution, is submitted and approved. This creates not only time but also legal constraints. It is also common for databases to become inaccessible after a key researcher leaves for another workplace, after a grant expires, legislation changes, after technical issues, and many other related reasons. In the field of iris recognition, we have found that despite the numerous freely available databases, these are often overlooked. One reason may be that a detailed overview of available databases is still missing. Existing review articles focus only on one selected area of iris recognition, and general surveys do not provide sufficient detail or scope to help guide this rapidly evolving and ever changing area. Based on the identified shortcomings, we have stipulated six research questions (RQ1-6) that we will try to answer.

RQ 1 What databases are used in the field of iris recognition and are publicly available to the scientific community ?

RQ 2 Which of these databases are the most popular ?

RQ 3 What are the common features of databases and how do they differ ?

RQ 4 What distinguishes popular databases and what are their common features ?

RQ 5 Which areas of iris recognition have few or no freely available databases ?

RQ 6 What are the recommendations for creating a new iris database ?

To answer these questions, we decided to create a list of available human iris datasets based on bibliometric research to perform statistical analysis that will help identify emerging directions and accelerate research in iris recognition and similar areas of research, as well.

4.1 Methods

Because there is a large number of scientific articles and other resources related to iris recognition, we have sought to select those that have the greatest impact on the scientific community in terms of the number of citations. We have chosen the Web of Science online library as an information source because it is publisher-independent and, to our knowledge, the most well-known works in the field of iris recognition are indexed there. In the Web of Science online library, we searched for keywords "recognition" and "iris", which had to contain the articles searched and restricted to English language only. We have further limited the search to Citation Index Expanded and to specific areas: IMAGING SCIENCE, OPTICS, TELECOMMUNICATIONS, PHOTOGRAPHIC TECHNOLOGY, MATHEMATICAL COMPUTATIONAL BIOLOGY, COMPUTER SCIENCE, AUTOMATION CONTROL SYSTEMS, MATHEMATICS, ENGINEERING, SCIENCE TECHNOLOGY, ROBOTICS, and OTHER TOPICS.

Based on this search, we obtained a list of 1,012 scientific publications in the field of research, which we then further analysed. Of all articles found, 689 relevant articles were manually selected based on text analysis and the authors of which used at least one iris database in the experimental section. We excluded from the analysis review studies, surveys, opinions, duplicates, and articles not related to iris recognition. An example is the article using the well-known Fisher's Iris data set that contains multivariate features of three species of Iris flowers [38, 39]. Despite these exclusions, not all databases could be directly assigned to a particular database or its version. If the analysed article used a preliminary version of a still expanding database (e.g., CASIA-IrisV1 - V4), we assigned it to the closest released version. In case we could not classify the database in any way due to lack of information, we considered it private. To verify the availability of databases, we decided to obtain all databases declared as publicly available with at least two email requests, and in some cases by phone, if necessary. Only under the assumption that we managed to gain access to the database, we marked it as publicly available. Finally,

we labelled all marginal cases according to the specific situation (e.g., no information available/no response, etc.).

4.2 Results

By a manual scan, we identified 158 datasets used in the field of iris recognition, most of which are declared publicly available. From all identified databases, we were able to obtain and verify access to 81 databases, which are publicly available for scientific purposes free of charge. In addition, we have identified 6 paid databases, 7 private databases, 30 datasets that are no longer available, and one database that is available for joint collaboration only. Due to the sharing of intellectual property rights, we did not include this database in the list of publicly available, but we created a separate group for it. We did not receive a response to 33 database requests, even after resubmitting and contacting all authors directly via email. A visual representation of the categorisation based on the availability of the iris databases is shown in Figure 2.

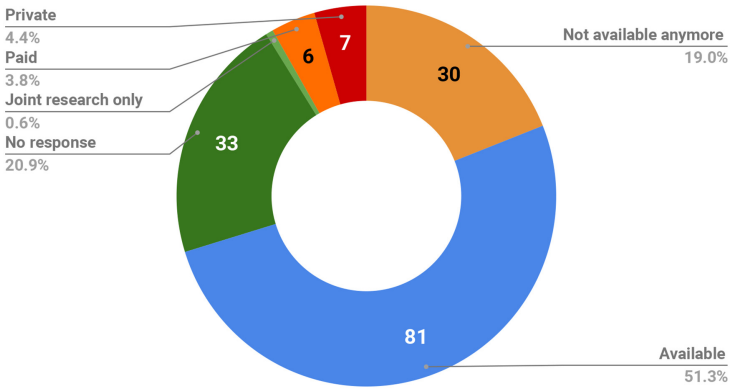

Figure 2: *The categorization of iris databases based on their availability.*

We further analysed the spectrum in which the individual databases were captured. The largest group consists of 102 databases captured in the near-infrared spectrum. The 39 databases captured in the visible spectrum form the second largest group and 14 databases were captured in both: near-infrared and visible spectrum simultaneously. The other group consists of 3 databases captured in the uncommon spectrum that do not fall into the previous categories, such as the thermal and near-ultraviolet spectrum (Figure 3). In terms of the total number of images, there are 1,257,468 images captured in the near-infrared spectrum and 121,089 images captured in the visible spectrum in all 81 available databases (1,378,867 iris images overall). In analysing the frequency of use of individual databases in the selected publications, we found that most authors use up to three databases, with the use of multiple datasets declining exponentially. Further

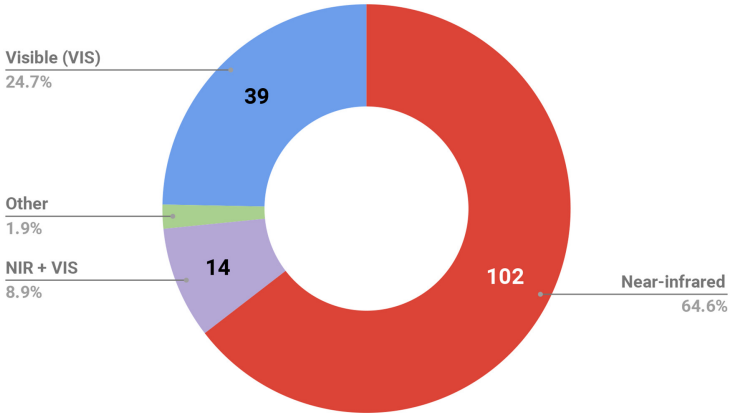

Figure 3: *The categorization of iris databases based on the spectrum in which they were captured.*

analysis showed that the multiple publicly available databases are created by only a few specialized research teams. The largest include, in no particular order: Center for Biometrics and Security Research of the Chinese Academy of Sciences (CASIA), Soft Computing and Image Analysis Group at the University of Beira Interior (SOCIA Lab), Biometrics and Machine Learning Group at the Warsaw University of Technology (ZBUM), Image Analysis and Biometrics Lab at Indian Institute of Technology Jodhpur (IAB), Biometric and Image Processing Lab at The University of Salerno via Ponte Don Melillo (BIPLab), Computer Vision Research Laboratory at The University of Notre Dame (CVRL), Biometrics and Identification Innovation Center at the West Virginia University, Biomedical Signal Analysis Lab at The Clarkson University (biosal), The National Science Foundation's' (NSF) Center for Identification Technology Research (CITeR), a multi-university center, Laboratory of Biometric Recognition and Information Security Technology at the Jilin University, Computer Vision and Pattern Recognition Unit at the Indian Statistical Institute, and The National Institute of Standards and Technology (NIST) at the U.S. Department of Commerce. Although many more research teams are involved in the field of iris recognition, most of them have created only one public database, or have decided to use a custom private database instead. This group was the most numerous and out of all 689 analysed publications, as much as 211 authors used the custom iris database in the experimental section. The most commonly used publicly available database is CASIA-V1 [40], although it is obsolete and its use is not recommended anymore [41]. Other very commonly used databases are Casia-IrisV3 [42, 43], UBIRIS-V1 [44] a UBIRIS-V2 [45]. The Bath (UBIID) database [46] is also very popular, although it is currently no longer available and is likely to be spread through internal sharing between researchers. Notable is also the increase in research interest in mobile databases in recent years, such as

MICHE-I [47], which indicates a change in the trend towards unconstrained iris recognition. For the identified 158 datasets, we collected basic 18 pieces of data, which were recorded in the form of a table. These include: the name of a database, information about availability, number of subjects, number of classes, number of images per subject, total number of images, resolution of images, year of creation, camera type, image format, image bit depth, used sensor, environment type, number of sessions, basic information about participants, link to the database website, brief description about the database and the original publication describing the dataset. Although this list has a great added value, it is obsolete by the definition, as new databases are created every month and the existing ones could disappear practically from day to day for various reasons. Therefore, to mitigate the effect of constant change, we have created an irisdata webpage (Available at: `https:\\irisdata.fei.stuba.sk`, and `http:\\irisdata.etrovub.be`) on which the entire list of identified datasets is available to the researchers and general public.

4.3 Discussion

In this chapter, we dealt with the analysis of existing human iris databases. Based on the bibiometric review, we manually created a list of 158 databases in order to find out the answer to 6 research questions (RQ 1-6). Based on efforts to verify individual databases by acquiring them, we found that 81 iris databases are currently available for scientific purposes, thus providing an answer to RQ 1. By statistical analysis of the use of individual databases, we confirmed the results of DeMarsico et al. [48] that the CASIA-V1 database is disproportionately more popular than other databases. In our analysis, the databases created and maintained by the Chinese Academy of Sciences (CASIA) are among the most used, while CASIA-V1 was referenced by 144 authors from among all 689 analyzed publications (answering RQ 2). For each database, we collected 18 pieces of data describing basic parameters and information about the contained images, and arranged them in tabular form. This allows to search, sort and classify diverse databases into specific groups as needed (answering RQ 3). In a deeper analysis of the most popular databases, we identified a few common characteristics that facilitate their extensive use by the scientific community: long-term accessibility, ease of access, database size, focus and timing of publication (answering RQ 4). Since the information websites of individual research institutions are still under development, there is often a contradiction between older published information in articles associated with created databases and current status, such as invalid or broken web links. Although these errors occur naturally, making efforts to ensure the stability of availability over the years by maintaining the same web address,

or redirecting the user if necessary, especially with regard to older databases, has a positive impact on its usability by various research teams. Another important factor is the restriction in terms of gaining access to the database files. Our analysis shows that the most popular databases have less strict conditions for obtaining access to the database for scientific purposes. They usually do not require anything other than an e-mail request (e.g. UPOL database [49]), and even if a license agreement is required, it can be signed directly by the requesting researcher (e.g. CASIA databases). However, this is not true for less used, albeit very high-quality databases, which usually require: a signature of the legal representative of the research institution (e.g. databases provided by The University of Notre Dame). a sharing of intellectual property rights (e.g. UBIPosePr database [50]), or fulfillment of a specific condition limiting access to the database (e.g. WVU Twins Day Dataset [51] is only available to researchers from the United States). Another common feature of the most used databases is a sufficiently large number of samples, which allow to perform statistically relevant and objective experiments. Last but not least, significant databases have brought new ideas and focus to previously unexplored areas of research, often making a given database the first in that field (e.g. UBIRIS-V1 [44] as the first database captured in the visible spectrum with the introduction of several noise factors). The created list of available databases naturally resulted in the identification of various areas that are dedicated to relatively non-standard forms of research and are not sufficiently represented by the available databases. These areas include, for example: analysis of the effects of various diseases on iris recognition, the effect of medications and drugs, the effect of lighting and other specific external factors, post-mortem iris recognition, and animal iris recognition. There is also a lack of databases dedicated to changing a single parameter under controlled conditions, which is the standard approach of scientific investigation. This often leads to inconclusive results that require further verification. To improve the current situation, it would help if the database creators provided metadata for the captured images (currently, their removal is more or less standard to ensure privacy) with detailed information not only about the sensors used, but also a precise description of the capture protocol and overall setup (answering RQ 5). Last but not least, recommendations for the creation of new iris databases can be directly drawn from common features of existing popular datasets (RQ 4). In addition to ensuring simple and reliable access to the database, providing a detailed description of the methodology and sensing protocols together with sensor metadata, and creating a sufficient number of samples for statistical evaluation, it is also necessary for the database to be unique and to help research new or underrepresented areas of the iris recognition (answering RQ 6).

CHAPTER 5

Achievements and Conclusions

In this work, we dealt with the use of neural networks in biomedicine with a specific focus on medical diagnostics. We have described the reasons for the introduction of artificial intelligence approaches in biomedicine and the possible forms of advancement of this newly emerging scientific field. Of all new methods, the most scientific attention is directed to deep learning approaches, which have seen tremendous progress not only in the field of medical diagnostics, but more or less in all scientific fields. Even though deep learning is being actively developed and is experiencing huge interest from the scientific community, there are still unsolved problems preventing the wider deployment of created deep neural models in everyday medical practice. Since during our studies we dealt with several areas of research; in this work we focused on the three most significant scientific contributions related to medical diagnostics that we managed to achieve and elaborated on them in more detail in individual chapters of this thesis.

The chapter 2: ***Improving Neural Network Models via Visual Priors***

- We extended the architecture of Receptive Fields Neural Networks inspired by scale-space theory. We modified the standard architecture using the Gaussian derivative basis so that it could use a general group of fixed filters with different shapes for the extraction of intrinsic features. We analysed the changes of fixed convolutional filters in shallow architectures with a limited number of training samples and evaluated the results in the form of a correlated Bayesian t-test. The results show that neural models with various tested bases achieve classification accuracy comparable to baseline and are very sensitive to normalization and sample selection. Due to the easier visualisation and interpretation of the extracted features and high accuracy in areas with insufficient data for the training of classical deep neural networks, these models have the potential to be used in everyday medical diagnostics.

The chapter 3: ***Localisation of Hard Exudates Based on Deep Learning***

- We proposed the use of a linear Support Vector Machine classifier to accelerate and increase localisation accuracy of hard exudates using a Faster Region-based Convolutional Neural Network. To increase the accuracy of object detector, we have introduced a filtering step using the feature vector extracted from the pre-trained ResNet-50 network in order to identify healthy samples and analyse potential findings in more detail. We evaluated the impact of the introduction of fast visual prescanning on the achieved localisation accuracy with and without contrast augmentation and interpreted the results using five-fold cross validation. The results show a significant reduction in false positive rate without a decrease in false negatives for entire images and for pixel-level evaluation. Due to the acceleration and improvement of visual detection of hard exudates in retinal fundus images, especially with paucity of training data, this method is suitable for incorporation into several possible clinical applications.

The chapter 4: ***Iris Imaging in Healthcare Setting***

- We provided a detailed overview of existing iris datasets created on the basis of a bibliometric analysis of relevant articles indexed in the Web of Science online library. On the basis of the popularity of their use in individual scientific articles, we have divided the analysed databases into categories corresponding to the subject area of research for which they are suitable. We also described popular datasets, their features and shortcomings, and provided recommendations for creating new databases for future research. We also briefly described the challenges and competitions related to iris recognition and non-standard forms of research in this area. To keep the created list up to date, the irisdata web domain has been created mainly due to the growing number of databases and rapid changes in availability. The created overview helps researchers identify suitable databases that can accelerate development in promising areas of research such as medical biometrics.

Bibliography

1. LISBOA, Paulo JG and IFEACHOR, Emmanuel C. *Artificial neural networks in biomedicine*. Springer Science & Business Media, 2000.
2. THOMPSON, Clive. Sure, A.I. Is Powerful - But Can We Make It Accountable? *Wired Magazine*. 2016, vol. 24, no. 10.
3. HUBEL, David H and WIESEL, Torsten N. Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*. 1959, vol. 148, no. 3, pp. 574–591.
4. ZEILER, Matthew D and FERGUS, Rob. Visualizing and understanding convolutional networks. In: *European conference on computer vision*. Springer, 2014, pp. 818–833.
5. OLAH, Chris, MORDVINTSEV, Alexander and SCHUBERT, Ludwig. Feature visualization. *Distill*. 2017, vol. 2, no. 11, e7.
6. LINDSAY, Grace W. Convolutional neural networks as a model of the visual system: Past, present, and future. *Journal of cognitive neuroscience*. 2021, vol. 33, no. 10, pp. 2017–2031.
7. CHOLLET, François. Xception: Deep learning with depthwise separable convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1251–1258.
8. WEI, Hang, WANG, Zulin and HUA, Gengxin. Dynamically Mixed Group Convolution to Lighten Convolution Operation. In: *2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD)*. IEEE, 2021, pp. 203–206.
9. JACOBSEN, Jorn-Henrik, VAN GEMERT, Jan, LOU, Zhongyu and SMEULDERS, Arnold WM. Structured receptive fields in cnns. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 2610–2619.
10. IMAMBI, Sagar, PRAKASH, Kolla Bhanu and KANAGACHIDAMBARESAN, GR. PyTorch. In: *Programming with TensorFlow*. Springer, 2021, pp. 87–104.
11. LECUN, Yann. The MNIST database of handwritten digits. *http://yann. lecun. com/exdb/mnist/*. 1998.
12. CRANE, Matt. Questionable answers in question answering research: Reproducibility and variability of published results. *Transactions of the Association for Computational Linguistics*. 2018, vol. 6, pp. 241–252.
13. DEMŠAR, Janez. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine learning research*. 2006, vol. 7, pp. 1–30.
14. RASCHKA, Sebastian. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808*. 2018.
15. BENAVOLI, Alessio, CORANI, Giorgio, DEMŠAR, Janez and ZAFFALON, Marco. Time for a change: a tutorial for comparing multiple classifiers through Bayesian analysis. *The Journal of Machine Learning Research*. 2017, vol. 18, no. 1, pp. 2653–2688.

16. CORANI, Giorgio, BENAVOLI, Alessio, DEMŠAR, Janez, MANGILI, Francesca and ZAFFALON, Marco. Statistical comparison of classifiers through Bayesian hierarchical modelling. *Machine Learning*. 2017, vol. 106, no. 11, pp. 1817–1837.

17. NILSSON, Adrian, SMITH, Simon, ULM, Gregor, GUSTAVSSON, Emil and JIRSTRAND, Mats. A performance evaluation of federated learning algorithms. In: *Proceedings of the second workshop on distributed infrastructures for deep learning*. 2018, pp. 1–8.

18. CORANI, Giorgio, BENAVOLI, Alessio, DEMŠAR, Janez, MANGILI, Francesca and ZAFFALON, Marco. *Baycomp* [`https://baycomp.readthedocs.io/en/latest/`]. 2017. Accessed: 2022-06-15.

19. SHAO, Jie, HU, Kai, WANG, Changhu, XUE, Xiangyang and RAJ, Bhiksha. Is normalization indispensable for training deep neural network? *Advances in Neural Information Processing Systems*. 2020, vol. 33, pp. 13434–13444.

20. HUANG, Lei, QIN, Jie, ZHOU, Yi, ZHU, Fan, LIU, Li and SHAO, Ling. Normalization techniques in training dnns: Methodology, analysis and application. *arXiv preprint arXiv:2009.12836*. 2020.

21. ULICNY, Matej, KRYLOV, Vladimir A and DAHYOT, Rozenn. Harmonic convolutional networks based on discrete cosine transform. *Pattern Recognition*. 2022, vol. 129, p. 108707.

22. ASHIR, Abubakar M, IBRAHIM, Salisu, ABDULGHANI, Mohammed, IBRAHIM, Abdullahi Abdu and ANWAR, Mohammed S. Diabetic Retinopathy Detection Using Local Extrema Quantized Haralick Features with Long Short-Term Memory Network. *International Journal of Biomedical Imaging*. 2021, vol. 2021.

23. NORIEGA, Alejandro, MEIZNER, Daniela, CAMACHO, Dalia, ENCISO, Jennifer, QUIROZ-MERCADO, Hugo, MORALES-CANTON, Virgilio, ALMAATOUQ, Abdullah, PENTLAND, Alex, et al. Screening Diabetic Retinopathy Using an Automated Retinal Image Analysis System in Independent and Assistive Use Cases in Mexico: Randomized Controlled Trial. *JMIR formative research*. 2021, vol. 5, no. 8, e25290.

24. LAKSHMINARAYANAN, Vasudevan, KHERADFALLAH, Hoda, SARKAR, Arya and JOTHI BALAJI, Janarthanam. Automated Detection and Diagnosis of Diabetic Retinopathy: A Comprehensive Survey. *Journal of Imaging*. 2021, vol. 7, no. 9, p. 165.

25. QURESHI, Imran, MA, Jun and ABBAS, Qaisar. Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning. *Multimedia Tools and Applications*. 2021, vol. 80, no. 8, pp. 11691–11721.

26. HE, Kaiming, ZHANG, Xiangyu, REN, Shaoqing and SUN, Jian. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

27. RUSSAKOVSKY, Olga, DENG, Jia, SU, Hao, KRAUSE, Jonathan, SATHEESH, Sanjeev, MA, Sean, HUANG, Zhiheng, KARPATHY, Andrej, KHOSLA, Aditya, BERNSTEIN, Michael, BERG, Alexander C. and FEI-FEI, Li. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*. 2015, vol. 115, no. 3, pp. 211–252. Available from DOI: 10.1007/s11263-015-0816-y.

28. REN, Shaoqing, HE, Kaiming, GIRSHICK, Ross and SUN, Jian. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*. Montreal, Canada: MIT Press, 2015, pp. 91–99. NIPS'15.

29. DECENCIÈRE, E., CAZUGUEL, G., ZHANG, X., THIBAULT, G., KLEIN, J.-C., MEYER, F., MARCOTEGUI, B., QUELLEC, G., LAMARD, M., DANNO, R., ELIE, D., MASSIN, P., VIKTOR, Z., ERGINAY, A., LAŸ, B. and CHABOUIS, A. TeleOphta: Machine learning and image processing methods for teleophthalmology. *IRBM*. 2013, vol. 34, no. 2, pp. 196–203. ISSN 1959-0318. Available from DOI: https://doi.org/10.1016/j.irbm.2013.01.010. Special issue : ANR TECSAN : Technologies for Health and Autonomy.

30. REDMON, Joseph and FARHADI, Ali. YOLO9000: better, faster, stronger. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 7263–7271.

31. REDMON, Joseph, DIVVALA, Santosh, GIRSHICK, Ross and FARHADI, Ali. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 779–788.

32. SÁNCHEZ, Clara I., NIEMEIJER, Meindert, IŠGUM, Ivana, DUMITRESCU, Alina, SUTTORP-SCHULTEN, Maria S.A., ABRÀMOFF, Michael D. and VAN GINNEKEN, Bram. Contextual computer-aided detection: Improving bright lesion detection in retinal images and coronary calcification identification in CT scans. *Medical Image Analysis*. 2012, vol. 16, no. 1, pp. 50–62. ISSN 1361-8415. Available from DOI: https://doi.org/10.1016/j.media.2011.05.004.

33. GIANCARDO, Luca, MERIAUDEAU, Fabrice, KARNOWSKI, Thomas P., LI, Yaqin, GARG, Seema, TOBIN, Kenneth W. and CHAUM, Edward. Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. *Medical Image Analysis*. 2012, vol. 16, no. 1, pp. 216–226. ISSN 1361-8415. Available from DOI: https://doi.org/10.1016/j.media.2011.07.004.

34. ZHANG, Xiwei, THIBAULT, Guillaume, DECENCIÈRE, Etienne, MARCOTEGUI, Beatriz, LAŸ, Bruno, DANNO, Ronan, CAZUGUEL, Guy, QUELLEC, Gwénolé, LAMARD, Mathieu, MASSIN, Pascale, CHABOUIS, Agnès, VICTOR, Zeynep and ERGINAY, Ali. Exudate detection in color retinal images for mass screening of diabetic retinopathy. *Medical Image Analysis*. 2014, vol. 18, no. 7, pp. 1026–1043. ISSN 1361-8415. Available from DOI: https://doi.org/10.1016/j.media.2014.05.004.

35. ABBASI-SURESHJANI, Samaneh, DASHTBOZORG, Behdad, HAAR ROMENY, Bart M. ter and FLEURET, François. Boosted Exudate Segmentation in Retinal Images Using Residual Nets. In: CARDOSO, M. Jorge, ARBEL, Tal, MELBOURNE, Andrew, BOGUNOVIC, Hrvoje, MOESKOPS, Pim, CHEN, Xinjian, SCHWARTZ, Ernst, GARVIN, Mona, ROBINSON, Emma, TRUCCO, Emanuele, EBNER, Michael, XU, Yanwu, MAKROPOULOS, Antonios, DESJARDIN, Adrien and VERCAUTEREN, Tom (eds.). *Fetal, Infant and Ophthalmic Medical Image Analysis*. Cham: Springer International Publishing, 2017, pp. 210–218. ISBN 978-3-319-67561-9.

36. KAUPPI, Tomi, KALESNYKIENE, Valentina, KÄMÄRÄINEN, J., LENSU, Lasse, SORRI, Iiris, RANINEN, A., VOUTILAINEN, R., UUSITALO, Hannu, KÄLVIÄINEN, Heikki and PIETILÄ, Juhani. The DIARETDB1 Diabetic Retinopathy Database and Evaluation Protocol. In: *BMVC*. 2007.

37. DECENCIÈRE, Etienne, ZHANG, Xiwei, CAZUGUEL, Guy, LAY, Bruno, COCHENER, Béatrice, TRONE, Caroline, GAIN, Philippe, ORDONEZ, Richard, MASSIN, Pascale, ERGINAY, Ali, CHARTON, Béatrice and KLEIN, Jean-Claude. FEEDBACK ON A PUBLICLY DISTRIBUTED IMAGE DATABASE: THE MESSIDOR DATABASE. *Image Analysis & Stereology*. 2014, vol. 33, no. 3, pp. 231–234. ISSN 1854-5165. Available from DOI: 10.5566/ias.1155.

38. FISHER, Ronald A. The use of multiple measurements in taxonomic problems. *Annals of eugenics*. 1936, vol. 7, no. 2, pp. 179–188.

39. DUA, Dheeru and GRAFF, Casey. *UCI Machine Learning Repository*. 2017. Available also from: http://archive.ics.uci.edu/ml.

40. MA, Lia, WANG, Yunhong, TAN, Tieniu, et al. Iris recognition based on multichannel Gabor filtering. In: *Proc. Fifth Asian Conf. Computer Vision*. 2002, vol. 1, pp. 279–283.

41. PHILLIPS, P Jonathon, BOWYER, Kevin W and FLYNN, Patrick J. Comments on the CASIA version 1.0 iris data set. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2007, vol. 29, no. 10, pp. 1869–1870.

42. MA, Li, TAN, Tieniu, WANG, Yunhong and ZHANG, Dexin. Personal identification based on iris texture analysis. *IEEE transactions on pattern analysis and machine intelligence*. 2003, vol. 25, no. 12, pp. 1519–1533.

43. WEI, Zhuoshi, TAN, Tieniu and SUN, Zhenan. Nonlinear iris deformation correction based on Gaussian model. In: *International Conference on Biometrics*. Springer, 2007, pp. 780–789.

44. PROENÇA, Hugo and ALEXANDRE, Luis A. UBIRIS: A noisy iris image database. In: *International Conference on Image Analysis and Processing*. Springer, 2005, pp. 970–977.

45. PROENÇA, Hugo, FILIPE, Silvio, SANTOS, Ricardo, OLIVEIRA, Joao and ALEXANDRE, Luis A. The UBIRIS. v2: A database of visible wavelength iris images captured on-the-move and at-a-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2009, vol. 32, no. 8, pp. 1529–1535.

46. MONRO, Donald M, RAKSHIT, Soumyadip and ZHANG, Dexin. DCT-based iris recognition. *IEEE transactions on pattern analysis and machine intelligence*. 2007, vol. 29, no. 4, pp. 586–595.

47. DE MARSICO, Maria, NAPPI, Michele, RICCIO, Daniel and WECHSLER, Harry. Mobile iris challenge evaluation (MICHE)-I, biometric iris dataset and protocols. *Pattern Recognition Letters*. 2015, vol. 57, pp. 17–23.

48. DE MARSICO, Maria, PETROSINO, Alfredo and RICCIARDI, Stefano. Iris recognition through machine learning techniques: A survey. *Pattern Recognition Letters*. 2016, vol. 82, pp. 106–115. ISSN 0167-8655. Available from DOI: https://doi.org/10.1016/j.patrec.2016.02.001. An insight on eye biometrics.

49. DOBEŠ, Mihal, MACHALA, Libor, TICHAVSKÝ, Petr and POSPÍŠIL, J. Human eye iris recognition using the mutual information. *Optik*. 2004, vol. 115, no. 9, pp. 399–404.

50. PADOLE, Chandrashekhar N and PROENÇA, Hugo. Compensating for pose and illumination in unconstrained periocular biometrics. *International Journal of Biometrics*. 2013, vol. 5, no. 3-4, pp. 336–359.

51. HOLLINGSWORTH, Karen, BOWYER, Kevin W and FLYNN, Patrick J. Similarity of iris texture between identical twins. In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*. IEEE, 2010, pp. 22–29.

Appendix A: List of Author‘s Publications

ADC

Scientific papers in foreign journals registered in Current Contents Connect

OMELINA, Ľ. - GOGA, J. - PAVLOVIČOVÁ, J. - ORAVEC, M. - JANSEN, B. A survey of iris datasets. In *Image and Vision Computing*, 2021, vol. 108, p. 104109. ISSN 0262-8856. Available online: https://doi.org/10.1016/j.imavis.2021.104109
(2021: 3.860 - IF, Q1 - SJR, Q1 - JCR (JIF), Q1 - JCR (JCI))
Number of citations (Google Scholar): 11

- ZANLORENSI, Luiz A., et al. Ocular recognition databases and competitions: A survey. Artificial Intelligence Review, 2022, 55.1: 129-180.
- HU, Junxing, et al. A large-scale database for less cooperative iris recognition. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2021. p. 1-6.
- MALGHEET, Jasem Rahman; MANSHOR, Noridayu Bt; AFFENDEY, Lilly Suriani. Iris recognition development techniques: a comprehensive review. In *Complexity*, 2021, 2021.
- KHADE, Smita, et al. Iris liveness detection for biometric authentication: A systematic literature review and future directions. In *Inventions*, 2021, 6.4: 65.
- ZAMBRANO, Jorge E., et al. Iris Recognition Using Low-Level CNN Layers Without Training and Single Matching. In *IEEE Access*, 2022, 10: 41276-41286.
- CHOUDHARY, Meenakshi; TIWARI, Vivek; VENKANNA, U. Identifying discriminatory feature-vectors for fusion-based iris liveness detection. In *Journal of Ambient Intelligence and Humanized Computing*, 2022, 1-12.
- SALEH, ABBADULLAH; SALEH, H. Iris Segmentation and Recognition Based On Deep Learning In The Presence Of Diseases. 2022. *PhD Thesis*.
- SAFA, Babak N.; BAHRANI FARD, Mohammad Reza; ETHIER, C. Ross. In vivo biomechanical assessment of iridial deformations and muscle contractions in human eyes. In *Journal of the Royal Society Interface*, 2022, 19.192: 20220108.
- TEMİZ, Hakan. Increasing classification accuracy through max voting method in iris recognition with support vector machine and hamming distance. In *Academic Research & Reviews In Engineering*.
- DAS, Sumanta; DE GHOSH, Ishita; CHATTOPADHYAY, Abir. Sclera biometrics in restricted and unrestricted environment with cross dataset evaluation. In *Displays*, 2022, 74: 102257.
- DE MARSICO, Maria, et al. Editorial to special issue on novel insights on ocular biometrics. In *Image Vis. Comput.*, 2021, 112: 104227.

KURILOVÁ, V. - GOGA, J. - ORAVEC, M. - PAVLOVIČOVÁ, J. - KAJAN, S. Support vector machine and deep-learning object detection for localisation of hard exudates. In *Scientific Reports*, 2021, vol. 11, p. 16045. ISSN 2045-2322. Available online: https://doi.org/10.1038/s41598-021-95519-0
(2021: 4.996 - IF, Q1 - SJR, Q2 - JCR (JIF), Q1 - JCR (JCI))
Number of citations (Google Scholar): 5

- THURZO, Andrej; KURILOVÁ, Veronika; VARGA, Ivan. Artificial Intelligence in Orthodontic Smart Application for Treatment Coaching and Its Impact on Clinical Performance of Patients Monitored with AI-TeleHealth System. In *Healthcare*. MDPI, 2021. p. 1695.
- WANG, Ting-Yuan, et al. Diabetic Macular Edema Detection Using End-to-End Deep Fusion Model and Anatomical Landmark Visualization on an Edge Computing Device. In *Frontiers in medicine*, 2022, 9.
- AI, Qing, et al. An improved MLTSVM using label-specific features with missing labels. In *Applied Intelligence*, 2022, 1-22.
- GOUTAM, Balla, et al. A Comprehensive Review of Deep Learning Strategies in Retinal Disease Diagnosis Using Fundus Images. In *IEEE Access*, 2022.
- BIBI, Nasira, et al. Automatic Detection of Exudates for Daignosis of Non-proliferative Diabetic Retinopathy using Region-based Convolutional Neural Networks. In *2021 International Conference on Frontiers of Information Technology (FIT)*. IEEE, 2021. p. 218-223.

ADAMÍK, M. - GOGA, J. - PAVLOVIČOVÁ, J. - BABINEC, A. - SEKAJ, I. Fast robotic pencil drawing based on image evolution by means of genetic algorithm. In *Robotics and Autonomous Systems*, 2022, vol. 148, p. 103912. ISSN 0921-8890. Available online: https://doi.org/10.1016/j.robot.2021.103912
(2021: 3.700 - IF, Q1 - SJR, Q2 - JCR (JIF), Q2 - JCR (JCI))
Number of citations (Google Scholar): 2

- SCALERA, Lorenzo, et al. Robotic Sponge and Watercolor Painting Based on Image-Processing and Contour-Filling Algorithms. In *Actuators*. MDPI, 2022. p. 62.
- WU, Pan-Long; HUNG, Yi-Chin; SHAW, Jin-Siang. Artistic robotic pencil sketching using closed-loop force control. In *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2022, 09544062221096946.

ADM

Scientific papers in foreign journals registered in Web of Science or Scopus

MÁCSIK, P. - PAVLOVIČOVÁ, J. - GOGA, J. - KAJAN, S. LBCNN for binary classification of diabetic retinopathy on fundus images. In *Acta Polytechnica Hungarica*, 2022, vol. 19, no. 7, p. 27-45, ISSN 1785-8860.
(2021: 1.711 - IF, Q2 - SJR, Q3 - JCR (JIF), Q1 - JCR (JCI))

ADF

Scientific papers in other domestic journals

KAJAN, S. - GOGA, J. Rozpoznávanie gest ruky pomocou konvolučných neurónových sietí. In *ATP Journal plus : Výskum v kybernetike na FEI STU v Bratislave*, 2019, vol. 2, p. 51-55. ISSN 1336-5010.

AEC

Scientific papers in foreign peer-reviewed proceedings, monographs

PIRHALA, M. - GOGA, J. - KURILOVÁ, V. - PAVLOVIČOVÁ, J. - KAJAN, S. Segmentation of significant regions in retinal images: perspective of U-Net network through a comparative approach. In *Systems, Signals and Image Processing. IWSSIP 2021. Communications in Computer and Information Science*, Springer, Cham. vol. 1527, p. 29-40. ISBN 978-3-030-96878-6. DOI: 10.1007/978-3-030-96878-6_3
(2021: Q4 - SJR)

AFC

Published papers from foreign scientific conferences

KAJAN, S. - GOGA, J. - LACKO, K. - PAVLOVIČOVÁ, J. Detection of diabetic retinopathy using pretrained deep neural networks. In *2020 Cybernetics & Informatics (K&I): 30th International Conference.* Velké Karlovice, Czech Republic. January 29-February 1, 2020. 1. ed. Danvers : IEEE, 2020, [5] p. ISBN 978-1-7281-4381-1. DOI: 10.1109/KI48306.2020.9039793.
Number of citations (Google Scholar): 8

- SUNGHEETHA, Akey; SHARMA, Rajesh. Design an early detection and classification for diabetic retinopathy by deep feature extraction based convolution neural network. In *Journal of Trends in Computer Science and Smart technology (TCSST)*, 2021, 3.02: 81-94.
- YAQOOB, Muhammad Kashif, et al. ResNet based deep features and random forest classifier for diabetic retinopathy detection. In *Sensors*, 2021, 21.11: 3883.
- JULIE, E. Golden; ROBINSON, Y. Harold; JAISAKTHI, S. M. (ed.). Handbook of Deep Learning in Biomedical Engineering and Health Informatics. CRC Press, 2021.
- SKARIAH, Shintu Mariam; ARUN, K. S. A Deep Learning Based Approach for Automated Diabetic Retinopathy Detection and Grading. In *2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE)*. IEEE, 2021. p. 1-6.
- TALLAPRAGADA, VV Satyanarayana. Deep Learning and Its Applications in Biomedical Image Processing. In *Handbook of Deep Learning in Biomedical Engineering and Health Informatics.* Apple Academic Press, 2021. p. 281-302.
- HERRERO TUDELA, María, et al. Processing of retinographs based on deep Learning to help diagnose diabetic retinopathy (in Spanish). 2020.
- MACSIK, Peter, et al. Local Binary CNN for Diabetic Retinopathy Classification on Fundus Images. In *Acta Polytechnica Hungarica*, 2022, 19.7.
- GEETHA, S., et al. Diabetic Retinopathy Grading with Deep Visual Attention Network. In *International Journal of Online & Biomedical Engineering*, 2022, 18.9.

KAJAN, S. - GOGA, J. - ZSÍROS, O. Comparison of algorithms for dynamic hand gesture recognition. In *2020 Cybernetics & Informatics (K&I): 30th International Conference*. Velké Karlovice, Czech Republic. January 29-February 1, 2020. 1. ed. Danvers : IEEE, 2020, [5] p. ISBN 978-1-7281-4381-1. DOI: 10.1109/KI48306.20-20.9039850. Number of citations (Google Scholar): 6

- HASTOMO, Widi. Gesture Recognition For Pencak Silat Tapak Suci Real-Time Animation. In *Jurnal Ilmu Komputer dan Informasi*, 2020, 13.2: 77-87.
- YUANYUAN, S. H. I., et al. Review of dynamic gesture recognition. In *Virtual Reality & Intelligent Hardware*, 2021, 3.3: 183-206.
- LABINTSEV, Andrey, et al. Recognition punches in karate using acceleration sensors and convolution neural networks. In *IEEE Access*, 2021, 9: 138106-138119.
- BHAGWAT, Aishwarya; GUPTA, Poonam; KADAM, Nivedita. Sign Language To Categorical Text using Convolutional Neural Network. In *2022 10th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing (ICETET-SIP-22)*. IEEE, 2022. p. 1-6.

- AMINUDIN, Widi Hastomo; RERE, LM Rasdi. Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information) 13/2 (2020), 77-87. DOI: http://dx: doi: org/10: 21609/jiki: v13i2. 855. In *Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)*, 2020, 13.2.
- FRANSLIN, Nurfazlin Muhamad Feizal; NG, Giap Weng. Vision-Based Dynamic Hand Gesture Recognition Techniques and Applications: A Review. In *Proceedings of the 8th International Conference on Computational Science and Technology*. Springer, Singapore, 2022. p. 125-138.

GOGA, J. - KAJAN, S. Hand gesture recognition using 3D sensors. In *Proceedings ELMAR-2017 : 59th International symposium*. Zadar, Croatia. 18-20 September, 2017. Zagreb : University of Zagreb, 2017, p. 181-184. ISSN 1334-2630. ISBN 978-953-184-230-3.

Number of citations (Google Scholar): 6

- YASEN, Mais; JUSOH, Shaidah. A systematic review on hand gesture recognition techniques, challenges and applications. In *PeerJ Computer Science*, 2019, 5: e218.
- PRAMUDITA, Aloysius Adya, et al. Contactless hand gesture sensor based on array of CW radar for human to machine interface. In *IEEE Sensors Journal*, 2021, 21.13: 15196-15208.
- LEE, Hunmin, et al. Real-time Interface Control with Motion Gesture Recognition based on Non-contact Capacitive Sensing. arXiv preprint arXiv:2201.01755, 2022.
- CHANDRASEKHAR, S.; MHALA, N. N. Performance Improvement of Data Fusion Based Real-Time Hand Gesture Recognition by Using 3-D Convolution Neural Networks With Kinect V2. In *Performance Improvement*, 2019, 9.1.
- WANG, Fei, et al. Gesture Recognition and Localization Using Convolutional Neural Network. In *2019 Chinese Control And Decision Conference (CCDC)*. IEEE, 2019. p. 5984-5989.
- VIGIL, M. S., et al. Indo-Pak Sign Language Translator Using Kinect. In *International Conference On Computational Vision and Bio Inspired Computing*. Springer, Cham, 2018. p. 725-732.

KAJAN, Slavomír - GOGA, Jozef. Simulation-based model control using static hand gestures in Matlab. In *Technical computing Prague 2017: 24th Annual conference proceedings*. Prague, Czech Republic. November 8, 2017. 1. ed. Prague : University of Chemistry and Technology, 2017, CD-ROM, [8] p. ISSN 2336-1662. ISBN 978-80-7592-002-7.

KAJAN, S. - PERNECKÝ, D. - GOGA, J. Application of neural network in medical diagnostics. In *Technical computing Prague 2015: 23r Annual conference proceedings.* Praha, 4.11. 2015. 1. vyd. Praha : University of Chemistry and Technology, 2015, CD-ROM, [9] p. ISSN 2336-1662. ISBN 978-80-7080-936-5.
Number of citations (Google Scholar): 5

- BILALOVIĆ, Omar; AVDAGIĆ, Zikrija. Robust breast cancer classification based on GA optimized ANN and ANFIS-voting structures. In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE, 2018. p. 0279-0284.
- HAGAN, Rachael; GILLAN, Charles J.; MALLETT, Fiona. Comparison of machine learning methods for the classification of cardiovascular disease. In *Informatics in Medicine Unlocked*, 2021, 24: 100606.
- RIAKINA, D. A.; LING, V. V. Artificial Neural Network For Medical Diagnostics. In *Mathematical methods and models in management, economics and sociology.* 2019. p. 179-184.
- NCHENA, Linos Mabvuto; JANÁČOVÁ, Dagmar. Fever status detection using artificial neuron network. In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021)*, Vol 1. Scitepress, 2021.
- RYAKINA, D. A.; BULASHOVA, A. N. Artificial Neural Network In Medical Field. In *Information and graphic technologies in professional and scientific activities.* 2018. p. 144-149.

Accepted for publication

KAJAN, S. - MACSIK, P. - KOROSI, L. - GOGA, J. - PAVLOVICOVÁ, J. Comparison of neural models for modeling dynamic changes in microgrids. *In 2022 Cybernetics & Informatics (K&I): 31th International Conference.* Visegrád, Hungary. September 11-14, 2022.

AFD

Published papers from domestic scientific conferences

KAJAN, S. - PAVLOVIČOVÁ, J. - GOGA, J. - KURILOVÁ, V. - KRAJŇÁK, R. - LACKO, K. Detection of optic disc and macula in retinal images using deep learning methods, In *Systems, Signals and Image Processing. IWSSIP 2021.*, Springer, Cham. Bratislava, Slovakia.

KURILOVÁ, V. - HUBA, M. - GOGA, J. - ORAVEC, M. - PAVLOVIČOVÁ, J. - MAJTÁNOVÁ, N. Two Machine-learning Approaches for Short-term COVID-19 Hospitalization Forecasting in Slovakia. In *Proceedings of the 21st Conference Information Technologies – Applications and Theory (ITAT 2021)*. Hotel Heľpa, Nízke Tatry and Muránska planina, Slovakia, September 24 – 28, 2021, p. 275-283. ISSN 1613-0073.

AGJ

Patent applications, utility model applications, design applications, trademark applications, applications for granting supplementary protection certificate

Pending - published application

SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA. Tool for redrawing a digital image by robotic drawing, method of redrawing an image by robotic drawing and method of tool calibration. Inventors: ADAMÍK, Michal and GOGA, Jozef. International Patent Classification: B25J 9/00, Utility model application. PUV 50091-2021. 24.11.2021. Slovakia.

GII

Other publications and documents

Unpublished - manuscript

GOGA, J. - VARGIC, R. - PAVLOVIČOVÁ, J. - ORAVEC, M. - KAJAN, S. Structure and base analysis of receptive fields neural networks in a character recognition task.