sciendo

DOI 10.2478/ama-2024-0054

acta mechanica et automatica, vol.18 no.3 (2024)
Special Issue "New Trends in Biomedical Engineering"

APPLICATION OF CONVOLUTIONAL GATED RECURRENT UNITS U-NET FOR DISTINGUISHING BETWEEN RETINITIS PIGMENTOSA AND CONE–ROD DYSTROPHY

Maria SKUBLEWSKA-PASZKOWSKA*, Pawel POWROZNIK*, Robert REJDAK, Katarzyna NOWOMIEJSKA****

*Faculty of Electrical Engineering and Computer Science, Department of Computer Science,
Lublin University of Technology, Nadbystrzycka 38D, 20-618 Lublin, Poland
**Faculty of Medicine, Chair and Department of General and Pediatric Ophthalmology,
Medical University of Lublin, Chmielna 1, 20-079, Lublin, Poland

maria.paszkowska@pollub.pl, p.powroznik@pollub.pl, robertrejdak@yahoo.com, katarzyna.nowomiejska@umlub.pl

received 22 July 2023, revised 19 January 2024, accepted 29 January 2024

Abstract: Artificial Intelligence (AI) has gained a prominent role in the medical industry. The rapid development of the computer science field has caused AI to become a meaningful part of modern healthcare. Image-based analysis involving neural networks is a very important part of eye diagnoses. In this study, a new approach using Convolutional Gated Recurrent Units (GRU) U-Net was proposed for the classifying healthy cases and cases with retinitis pigmentosa (RP) and cone–rod dystrophy (CORD). The basis for the classification was the location of pigmentary changes within the retina and fundus autofluorescence (FAF) pattern, as the posterior pole or the periphery of the retina may be affected. The dataset, gathered in the Chair and Department of General and Pediatric Ophthalmology of Medical University in Lublin, consisted of 230 ultra-widefield pseudocolour (UWFP) and ultra-widefield FAF images, obtained using the Optos 200TX device (Optos PLC). The data were divided into three categories: healthy subjects (50 images), patients with CORD (48 images) and patients with RP (132 images). For applying deep learning classification, which rely on a large amount of data, the dataset was artificially enlarged using augmentation involving image manipulations. The final dataset contained 744 images. The proposed Convolutional GRU U-Net network was evaluated taking account of the following measures: accuracy, precision, sensitivity, specificity and F1. The proposed tool achieved high accuracy in a range of 91.00%–97.90%. The developed solution has a great potential in RP diagnoses as a supporting tool.

Key words: retinitis pigmentosa, convolutional GRU U-Net, classification, UWFP, UWFAF, deep learning

1. INTRODUCTION

Retinitis pigmentosa (RP) is a group of inherited retinal diseases characterised by the progressive dysfunction of rod and cone photoreceptors in the retina. The majority of cases are inherited in Mendelian patterns, namely autosomal dominant (30%–40% of cases), autosomal recessive (50%–60%) or X-linked (5%–15%) inheritance. Depending on the type of cell primarily affected, RP can be categorised into rod-dominant (classical RP) and cone-dominant cone–rod dystrophy (CORD) [5]. According to RetNet [22], more than 300 causative genes have been identified for RP so far, of which 58 genes have been identified in families with autosomal recessive RP. It is the most common inherited retinal dystrophy globally, affecting 1 in 4,000 individuals [8, 13]. In the past two decades, extensive genetic studies on RP not only led to the identification of the molecular basis in at least 60% of families but also set the basis for gene-based therapy. Although new treatments for RP such as gene therapy are being developed [25], current practice mainly involves care for residual visual function and surgery or medical therapy for complications. Thus, an appropriate clinical evaluation and estimation method for residual visual function and structure in patients with RP should be established. Identifying potential anatomical biomarkers for disease progression in RP is highly relevant for assessing treatment end points in RP in clinical trials.

Artificial Intelligence (AI) has gained a prominent role in the medical industry. The rapid development of the computer science field has caused that AI to become a meaningful part of modern healthcare. Image-based analysis involving neural networks is a very important part of eye diagnoses [18]. These modern approaches allow us to find and localise the patterns that are characteristic of the particular disease. The neural network methods are the perfect tool to distinguish between healthy people and patients suffering from various disorders. Moreover, not only particular diseases but also their place of occurrence may be recognised.

The methods and system involving AI are additional support indicating the proper diseases or alerting when the specific features are located [18, 19, 21]. These methods may increase the detection of diseases, especially rare diseases, such as retina problems, in the areas where there are no specialists. The patients would gain the knowledge from AI. Moreover, AI will definitely shorten the consultation time.

Applying AI algorithms reduces the diagnosis time, supports decision-making and avoids misdiagnosis, and reduces the cost of treatment or provides the treatment recommendation [16].

Despite the undoubted advantages of existing solutions, there is still a niche in the quick and effective diagnosis of eye diseases, especially rare or genetic diseases. The main motivation of this study is to apply AI for the classification of eye changes caused by RP disease. It is very important to distinguish between RP and CORD.

The main aim of this study is to classify RP. The Convolutional Gated Recurrent Units (GRU) U-Net was created for this purpose. The dataset, gathered in the Chair and Department of General and Pediatric Ophthalmology of Medical University in Lublin, consisted of 230 ultra-widefield fundus photography (UWFP) and ultra-widefield FAF (UWFAF) images. For applying deep learning classification, which rely on a large amount of data, the dataset was artificially enlarged using augmentation involving image manipulations. Three classes were defined: healthy cases, cases with CORD and cases with RP on the periphery of an eye.

This study proposes that Convolutional GRU U-Net is an adequate tool for the classification of retina dystrophy and healthy cases.

The following research questions have been formulated:

- RQ1: Is it possible to achieve satisfactory classifier accuracy for the proposed the Convolutional GRU U-Net model?
- RQ2: Is it possible to distinguish RP from CORD and healthy cases from Optos images?
- RQ3. Is it possible to achieve satisfactory classifier accuracy while reducing the raining time?

2. RELATED WORKS

A great number of studies concern application of AI for medical purposes, including U-Net solutions.

2.1. RP

Patients with RP typically experience night blindness and progressive loss of peripheral vision in both eyes during their first decades of life [36]. This leads to the loss of central vision and progression to total blindness. Typical form of RP is a long-lasting disease that usually evolves over several decades. However, there are extreme cases with a rapid evolution over two decades or a slow progression that never leads to blindness. This long-lasting disease is classified into three stages: early, middle and end stages. In the early stage, night blindness is the main symptom. In the mid stage, visual field testing shows mild periphery scotomas that tend to enlarge towards extreme periphery and macular areas. In the end stage, patients can no longer move autonomously, as a result of peripheral vision loss (classical tunnel vision), with few degrees of remaining visual field around this the main diagnostic tool to assess the visual function.

Variable expression in fundus changes may be present in different patients with different causes, in different patients caused by mutations in the same gene, in different patients from the same family and probably even in different eyes of the same patient.

The classic ophthalmoscopic findings in RP are typically described as a triad: retinal blood vessel attenuation, waxy pallor of the optic disc and intraretinal pigmentation in a bone–spicule pattern in the periphery of the retina. Abnormal retinal pigmentation occurs when pigment from disintegrating retinal pigment epithelial cells migrates into the superficial (“inner”) retina as a response to photoreceptor cell death [22]. Initially, the pigmentary abnormality appears as a fine dusting extending from the mid to the far peripheral retina. Later, “bone-spicules” form throughout the mid and the far retinal periphery due to accumulations of pigment surrounding retinal blood vessels. In advanced RP, atrophy of the choriocapillaris may expose the large choroidal vessels beneath. CORD presents with visual acuity loss, photophobia, and dyschromatopsia, showing abnormal cone photopic electroretinography (ERG) response with no or mild rod involvement and pigmentary changes within the posterior pole [12].

The diagnosis and monitoring of RP require a comprehensive approach, encompassing patient history, fundus examination with fundus wide-field photos, wide-field fundus autofluorescence (FAF), visual field and full-field ERG. These diagnostic tools play a crucial role in detecting early signs of RP in at-risk individuals and monitoring the progression of the disease.

There is a progress towards preventing the loss of or restoring the function of rod photoreceptors in retinal dystrophies. There is also a need for a standardised device for quantifying the loss of the structure and function of the retina in RP. The emergence of molecular therapies for inherited retinal degenerations has highlighted the need for imaging modalities that can sensitively identify structural disease progression and robust methods to analyse disease progression.

Both UWFP and UWFAF imaging using a scanning laser ophthalmoscope are imaging techniques that enable clinicians to obtain fundus images with a 200° angle of view easily and noninvasively [24]. FAF is a noninvasive method that evaluates retinal disease and reflects retinal pigment epithelium functions by visualising the accumulation of lipofuscin. FAF images of patients with RP show hyperfluorescence in the early disease stages, whereas hypofluorescence corresponds to lesions in later stages. In typical RP, an autofluorescence ring, which represents a hyperfluorescent ring in FAF images, might be observed at the border separating the dysfunctional retina from the functional retina [27, 30].

2.2. Tools for RP classification

Various appliances for improving the modern eye healthcare have been developed using AI. The new models have been created to classify various eye diseases, including RP, as a tool supporting diagnoses. Modern systems have also been developed for these purposes. Many types of images have been taken into consideration.

The identification of normal and selected diseases based on ultra-wide field imaging is described in a previous study [35]. Three deep learning networks, namely EfficientNet-B7, DenseNet and ResNet-101 were applied to classify diabetic retinopathy, retinal vein occlusion, pathologic myopia, retinal detachment, RP, age-related macular degeneration, vitreous opacity and optic neuropathy. The best accuracy was obtained for the EfficientNet-B7 network.

Image-based diagnosis using the Deep Neural Network (DNN) model Visual geometry group-16 (VGG-16) has been applied to classify patients suffering from RP disease and healthy patients [21]. Two types of images were involved: 373 UWFP and 373 UWFAF. The obtained results proved that the proposed model is efficient for RP recognition (UWFP: sensitivity 99.3%, specificity 99.1%, and UWFAF: sensitivity 100%, specificity 99.5%). It was also stated that further studies involving AI are required for RP diagnosis.

Glaucoma, maculopathy, pathological myopia and RP were classified using three deep learning approaches, namely MobileNetV2, InceptionV3 and AlexNet, in a previous study [14]. Due to small amount of data, transfer learning with ImageNet was applied. The Kaggle dataset was used for the studies. The best results obtained MobileNetV2: up to 98.4% for accuracy, 96% for sensitivity and 99% for specificity.

A previous study has used Convolutional Neural Network (CNN) for automatic recognition of healthy people and people with the diseased retina based on retinal fundus images [18]. The images were obtained from the Friedrich-Alexander University machine learning data repository and from the eye hospital in Bangalore, India. Both data were augmented to obtain the proper number of images for classification. The average accuracy was reached between 96.5 % and 99.7%.

Four pretrained neural networks, namely AlexNet, DenseNet-161, ResNet-50 and ResNet-152, were applied for infrared (IR) images, optical coherence tomography (OCT) images and the two combined [20]. The highest precision was obtained for the combination of IR and OCT images. The studies were performed on the Johns Hopkins University (JHU) dataset.

The end-to-end deep learning approach consisting of DenseNet, Recurrent Neural Network and Fully Connected layer was applied for predicting treatment requirements for the management of neovascular age-related macular degeneration [28]. Three classes were defined as low, intermediate and high. The obtained accuracy was up to 72%, sensitivity was 82% and specificity was 71%.

The RP classification based on OCT images using pixel-based and component-based methods was described in a previous study [8]. Four classifiers, namely Random Forest pixel-based, AdaBoost.M1 pixel-based, Random Forest component-based and AdaBoost.M1 component-based, were applied with great success. The obtained accuracy exceeded 99%. The precision was in the range of 42.45%–68.40%, the recall between 52.2% and 79.4%, and F1 between 46.76% and 64.69%.

Classification methods were also applied for developing expert systems for various eye disease. An expert system, dedicated to diabetic retinopathy and RP, consisting of extracting colour channels, red and green, was used [10]. Then, the Sobel method for edge detection was applied for the red channel, and the region of interest was determined for the green channel. In the next step, pixels with high intensity from the red channel and low-intensity pixels from the green channel were merged. Finally, preprocessing included the masking and removal of unwanted regions, and pixels were interpolated to the original image. The data were obtained using a CANON CF-60UVi camera with a 60o FOV. The achieved accuracy was 85.45%.

Another expert system for RP classification was used in a previous study [7]. It aimed at the automatic segmentation of pigment deposits in the fundus of the eye images. OCT images and those obtained from fundus camera were preprocessed. Then, the watershed transformation was applied to obtain homogeneous regions. This made the proposed method highly resistant to high variabilities in pigment deposits in terms of colour and shape. Even small pigment deposits were recognised. The system accuracy, precision, recall and F1 score were 97.90%, 74.43%, 98.44% and 59.04%, respectively.

After analysing the studies concerning applying AI for RP diagnosis, the authors noticed that there are no studies involving the location of retinal changes based on Optos images. Therefore, the aim of this study is to propose a Convolutional GRU U-NET classifier to distinguish between a healthy eye and an eye with pigmentary changes located in the centre (CORD) or on the periphery of the retina (RP).

3. MATERIALS AND METHODS

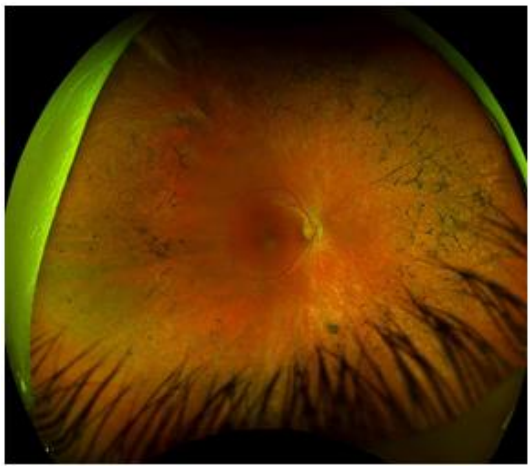
3.1. Dataset

UWFP and UWFAF images have been obtained using an Optos 200TX device (Optos PLC). Medical records of patients with cone–rod and cone–rod dystrophies at the Chair and Department of General and Pediatric Ophthalmology of the Medical University of Lublin were retrospectively reviewed. Patients were included if they had UWFAF and clinical characterisation of their retinal dystrophy. UWFP and UWFAF imaging were performed after pupil dilation with topical 0.5% tropicamide. All patients underwent visual acuity testing, slit lamp biomicroscopy and dilated fundus examination. All patients had clinical findings consistent with RP or CORD. UW-FAF characteristics were analysed qualitatively by two reviewers. The study included 69 eyes of 35 patients. All patients were identified and diagnosed clinically by an inherited retinal disease specialist. The control group consisted of 15 healthy subjects (30 eyes). A total of 230 optomap retinal images were obtained. The data were divided into three categories: healthy subjects (50 images), patients with CORD (48 images) and patients with RP on the periphery of the retina (132 images). For applying deep learning classification, which requires a large amount of data, the dataset was artificially enlarged using augmentation involving image manipulations [38]. The images were rotated by 10°. Additionally, they were converted to greyscale. Previous study involving CNN and their various modifications, has proven how important role plays a properly selected and sufficiently large dataset. However, in studies using medical data, there is almost always a problem of shortage of training data due to various difficulties in obtaining data. Data augmentation is commonly used in such cases. As numerous studies indicate [1,2,6,32,37] that data augmentation improves the quality of classification, regardless of the measure used. In this study, the augmented dataset was exclusively utilised for training the model.

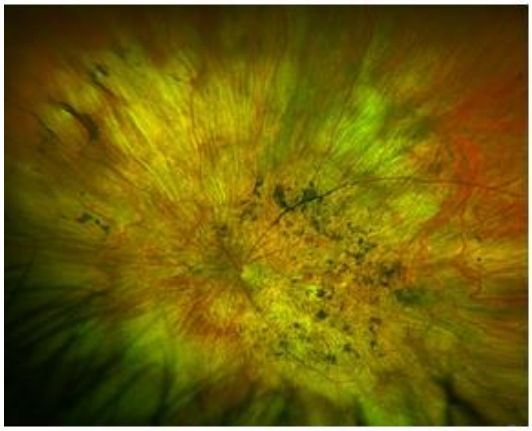
Deep learning models were employed for acquiring RP features and associated classifications from the complete input images. Automatically, irrelevant regions, including those beyond the circular boundary of the main object, were eliminated. These unimportant areas may arise from artefacts introduced during image acquisition. Subsequent to the removal of irrelevant portions, the images were resized to 300 × 375 pixels, with pixel values ranging from 0 to 1.

The final dataset, used in this study, contained the following number of images: 264 for healthy persons, 240 for patients with RP in the centre of the retina and 240 for patients with RP on the periphery of the retina. The images representing RP and CORD are shown in Fig. 1. In case of RP, the most typical changes in the fundus of the eye are narrow vessels and pigmented changes in the periphery of the fundus. The macular area is characterised by shiny reflexes. In case of CORD, the changes are located in the centre of the retina.

The final dataset, used in this study, contained the following number of images: 264 for healthy persons, 240 for patients with RP in the centre of the retina and 240 for patients with RP on the periphery of the retina. The images representing RP and CORD are shown in Fig. 1. In case of RP, the most typical changes in the fundus of the eye are narrow vessels and pigmented changes in the periphery of the fundus. The macular area is characterised by shiny reflexes. In case of CORD, the changes are located in the centre of the retina.



RP



CORD

Fig. 1. Example images from dataset representing RP and CORD dystrophy. Wide-field fundus photography obtained using Optos device

3.2. GRU

The GRU model, widely recognised as a prominent variation of Long Short-Term Memory (LSTM) [13], combines the forgetting gate and input gate into a single update gate while also merging the cell state and hidden state. Consequently, the resulting GRU model is simpler and faster than the conventional LSTM model. This characteristic proves particularly advantageous when handling large datasets, allowing for substantial time savings with only minimal performance discrepancies compared to the standard LSTM model. Both LSTM and GRU models excel at preserving crucial features by using various gates, ensuring these features remain intact even during long-term transmissions. At time t, the new state can be calculated using Eq. (1) [9]:

$$u_t = (1 - z_t) \circ u_{t-1} + a_g \circ \widetilde{u_t} \quad (1)$$

where u_t is the new GRU state, u_{t-1} is the previous GRU state, a_g is the update gate and $\widetilde{u_t}$ is the current candidate sate with a new collection of information.

The update gate plays a crucial role in determining the balance between retaining past information and incorporating new information. Its primary function is to regulate the extent to which information from the previous state influences the current state. By adjusting the value of a_g, the degree to which the information from the previous state is integrated into the current state is controlled. A higher value of a_g signifies a greater incorporation of information from the previous state. Eq. (2) defines the process of update gate modification [9]:

$$a_g = \delta(W_z x_t + R_z u_{t-1} + b_z) \quad (2)$$

where x_t is the data vector at time t, W_z is update gate weights at time t, R_z is update gate weights at time $t-1$ and b_z is bias.

The current candidate state is calculated using Eq. (3) [9]:

$$\widetilde{u_t} = \tanh(W_h x_t + r_t \circ R_z u_{t-1})\, b_z \quad (3)$$

where r_t is a reset gate at time t. Its role is to control the flow of information between current and previous states.

In numerous modelling tasks, having access to both past and future states proves advantageous. However, Conventional GRU networks process sequences in a temporal order, disregarding any future state. To address this limitation, Bidirectional GRU networks are introduced, which expand unidirectional GRU networks by incorporating a second layer. In this additional layer, the hidden-to-hidden connections flow in the opposite temporal order, enabling the model to capture the future state alongside the traditional past state. This bidirectional approach enhances the network's ability to understand and leverage information from both temporal directions.

3.3. Proposed methodology

This network is inspired from previous studies [3, 15, 29, 34]. The network utilises the strengths of Bidirectional GRU networks as well as densely connected convolutions. The individual components are described in detail in the following sections. The whole structure of it is presented in Fig. 2.

3.3.1. Encoding

The modified U-Net's contracting path involves a series of four steps. Each step comprises two convolutional filters of a size of 3×3, followed by a 2×2 max pooling function and the Rectified Linear Unit (ReLU) activation. At each step, the number of feature maps is doubled. The contracting path progressively extracts image representations while increasing the dimensionality of these representations layer by layer. Ultimately, the final layer in the encoding path produces a high-dimensional image representation with rich semantic information.

In the original U-Net, the last step of the encoding path consists of a sequence of convolutional layers. This enables the network to learn diverse types of features. However, this approach may result in the learning of redundant features through successive convolutions. To address this issue, densely connected convolutions were introduced [15]. This enhancement allows the network to improve its performance by leveraging the concept of "collective knowledge", which involves concatenating feature maps learned from all previous convolutional layers with the feature map obtained from the current layer. These concatenated feature maps are then forwarded as input to the subsequent convolutional layer. This strategy promotes the reuse of valuable feature maps throughout the network, mitigating the risk of learning redundant features.

The idea of densely connected convolutions has some advantages over regular convolutions [15]. First of all, it helps the network to learn a diverse set of feature maps, instead of redun-

dant features. Moreover, this idea improves the network's representational power by allowing information flow through the network and reusing features. Furthermore, densely connected convolutions can benefit from all the produced features before it, which prompts the network to avoid the risk of exploding or vanishing gradients. In addition, the gradients are sent to their respective places in the network more quickly in the backward path. The idea of densely connected convolutions was implemented in the proposed network, due to which two consecutive convolutions are introduced as one block. There is a sequence of N blocks in the last convolutional layer of the encoding path, as shown in Fig. 3. These blocks are densely connected.

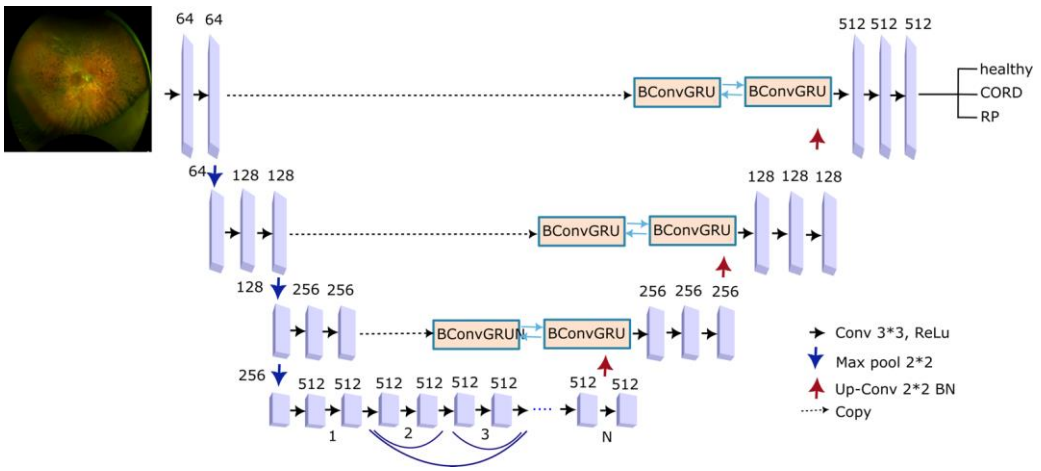


Fig. 2. The architecture of Convolutional GRU U-Net segment-based

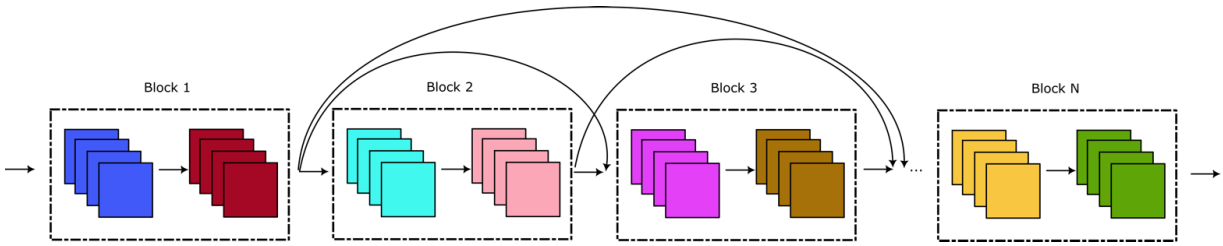


Fig. 3. U-Net dense layer

3.3.2. Decoding

The decoding path begins by applying an up-sampling function to the output of the previous layer. In the traditional U-Net model, the corresponding feature maps from the contracting path are cropped and copied to the decoding path. These feature maps are then concatenated with the result of the up-sampling function. However, in this approach, the Bidirectional Convolutional GRU was utilised to process these two types of feature maps in a more intricate manner (Fig. 4).

The set of feature maps copied from the encoding part is initially passed through an up-convolutional layer. This layer applies an up-sampling function, followed by a 2×2 convolution operation, effectively doubling the size of each feature map and halving the number of feature channels. Consequently, the resulting feature maps progressively increase in size throughout the expanding path, layer by layer, until they reach the original size of the input image after the final layer.

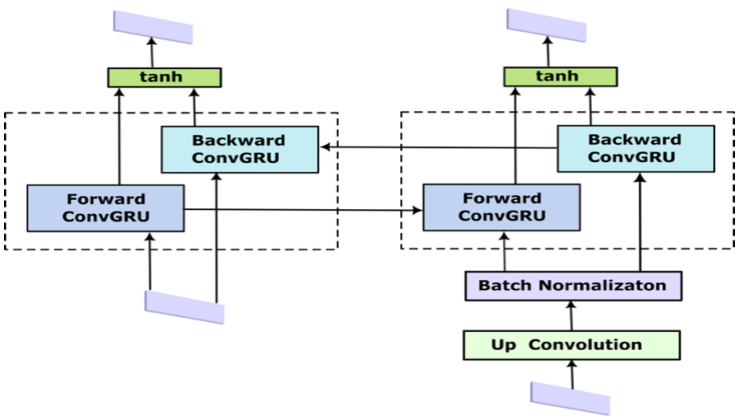


Fig. 4. Bidirectional ConvGRU

3.3.3. Batch normalisation

Following the up-sampling, the feature map is processed by a batch normalisation (BN) operation. During the training phase, intermediate layers often encounter a challenge where the distribution of activations tends to vary. This issue can significantly slow down the training process as each layer needs to adapt to new distributions in every training step. To address this problem, BN [17] is employed to enhance the stability of the neural network. BN standardises the inputs to a layer by subtracting the batch mean and dividing it by the batch standard deviation (SD). This normalisation process effectively accelerates the training speed of the neural network. Additionally, in certain cases, BN provides a regularisation effect, which can further improve the model's performance.

3.3.4. Bidirectional Convolutional GRU

The output of the BN step is fed to the Bidirectional Convolutional GRU layer. The main disadvantage of the standard GRU is that these networks does not take account of the spatial correlation since these models use full connections in input-to-state and state-to-state transitions. To solve this problem, bidirectional convolutional GRU was proposed, which exploited convolution operations. The idea of this layer based on LSTM was presented in a previous study [31]. It consists of an input vector $\mathrm{x_t}$, an output vector $\mathrm{u_t}$, a reset gate $\mathrm{r_t}$, update gate $\mathrm{a_g}$ and candidate activation vector $\widetilde{\mathrm{u_t}}$. The update rule for the input vector $\mathrm{x_t}$ and the previous output $\mathrm{u_{t-1}}$ is given by the following equations (for convenience, the subscript and superscript from the parameters were removed):

$$r = \sigma(W_{r*n}[u_{t-1}; x_t] + b_r) \quad (4)$$

$$a = \sigma(W_{u*n}[u_{t-1}; x_t] + b_u) \quad (5)$$

$$c = \rho(W_{c*n}[x_t; r \odot u_{t-1}] + b_c) \quad (6)$$

$$u_t = \mathrm{a} \odot u_{t-1} + (1 - \mathrm{a}) \odot c \quad (7)$$

where σ and ρ are sigmoidal and ReLU functions, respectively; $*\,\mathrm{n}$ denotes a convolution kernel of size $\mathrm{n} \times \mathrm{n}$; $\odot$ indicates Hadamard product operation. Brackets are presented to indicate a feature concatenation; and $\mathrm{r, a, c}$ and $\mathrm{u_t}$ denote the typical elements (reset, update, current memory content and a new state) of GRU.

4. RESULTS

4.1. Classifier evaluation

The proposed Convolutional GRU U-Net segmentation-based network was evaluated taking account of the following measures [4]: accuracy (Eq. 8), precision (Eq. 9), sensitivity (Eq. 10), specificity (Eq. 11) and F1 score (Eq. 12). A series of experiments were performed considering a random split of the data into training and testing sets: 80% and 20%, respectively. To reduce the randomness of the results, 10 iterations were performed independently. To assess the effectiveness of the classifier, commonly used measures were used: accuracy, precision, sensitivity, specificity and the F1 score [23,33].

Tab. 1. Accuracy results obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
All	94.05	97.90	91.00	3.97

Tab. 2. Accuracy results of healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	95.48	97.91	92.14	3.86
CORD	94.87	97.98	92.07	4.89
RP	95.15	97.92	92.09	2.90

The accuracy results for the proposed Convolutional GRU U-Net for RP classification based on Optus images are provided in Tab. 1. This measure gives the information about the ability to distinguish between healthy cases and the cases with RP disease. The obtained values confirm that the segmentation method involving deep learning is a suitable tool for RP recognition purposes. The mean accuracy exceeds 94%, which proves that the classification RP has been developed with great success. The detailed accuracy results for three defined classes are presented in Tab 2. The proposed tool recognises heathy eyes, cases CORD and pigmentary changes localised on the periphery of the retina. The obtained minimum accuracy was greater than 92%. The maximum measures almost reached 98%.

Precision results for the developed tool is presented in Tab. 3. The mean value exceeds 95% for each class and ranges between 90% and 98%. The high value of this measure means that the network does many correct classifications or only few misclassifications.

Tab. 3. Precision results of healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	95.85	98.36	91.90	3.74
CORD	95.47	98.36	91.90	4.70
RP	95.32	98.41	90.00	3.90

Tab. 4. Sensitivity results of healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	94.71	97.56	91.54	3.47
CORD	93.41	96.82	90.27	3.82
RP	95.11	97.91	91.38	2.75

Sensitivity measures how correctly the model can classify positive instances. The obtained results, presented in Tab. 4, indicate that the developed tool has a great performance. The mean sensitivity exceeded 93%, while the values were in the range of 90.27%–97.91%.

Specificity measures the amount of real negative data that are classified as negative ones. The obtained values are summarised in Tab. 5. As it can be seen, the mean measure exceeds 92%, while the whole calculated indicator is between 89.64% and 99.17%. The higher specificity value was obtained for cases with RP diseases located in the centre of the retina.

Tab. 5. Specificity results of healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	95.42	98.36	92.28	3.61
CORD	98.46	99.17	95.36	3.27
RP	92.94	97.64	89.64	2.50

The harmonic mean of F1 scores was calculated based on precision and specificity. Based on the previous values of precision and specificity, it can be assumed that this measure also gained high values. The mean value exceeded 94%. The proposed tool recognises the cases with RP located in the central of the retina the best. Slightly worse results were obtained for the cases with RP located in the periphery of the retina and healthy ones. However, these differences are very small, up to 2.79%. The SD in all measures (Tabs. 1–6) is very low, which means that the observations are gathered around the mean and thus repeatable.

The learning parameters of the proposed tool are presented in Figs. 5 and 6. As can be seen from Fig. 5, after reaching 40 epochs, the accuracy, both for training and test sets, is on the same level (high). Another element indicating how well the neural network works is the loss value. As can be observed in Fig. 6, 50 epochs are enough to minimise the loss. These two parameters prove that the learning process was enough to obtain high accuracy values.

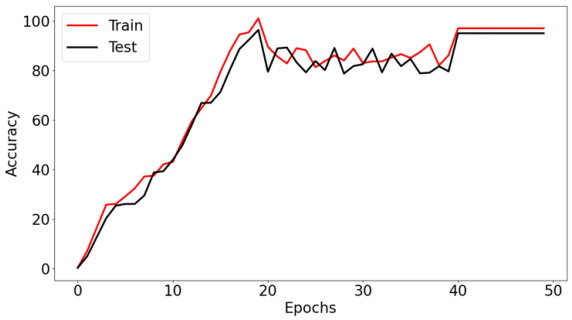


Fig. 5. Accuracy obtained for training and test data within 50 epochs. At least 40 epochs are required to obtain the accuracy exceeding 90%

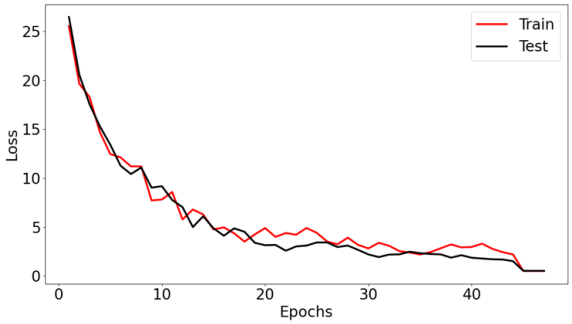


Fig. 6. Values of loss function while training and testing the model. It can be observed that almost 50 epochs are required to achieve the acceptable error level

The confusion matrix, depicted in Fig. 7, presents the classes that were misclassified. As it can be seen, two types of RP disease are confused, as well as the RP cases with healthy cases. The level of misclassification is very small, up to 5.14%. This matrix also is also suitable for the neural network. To ensure the accuracy of the developed model, the Leave-One-Out Cross-Validation (LOOCV) was performed. Although this procedure is computationally intensive, it provides reliable and unbiased insights into the model's performance. By employing LOOCV, the root mean squared error (RMSE) for n tests was calculated. The obtained results for LOOCV were as follows: RMSE: 5.83% and SD ±3.78%.

Tab. 6. F1 score results for healthy, RP and CORD cases obtained using Convolutional GRU U-NET

Class	Mean (%)	Max (%)	Min (%)	±SD (%)
Healthy	95.63	98.36	92.10	3.65
CORD	96.94	98.76	93.99	3.35
RP	94.11	98.02	89.82	3.11

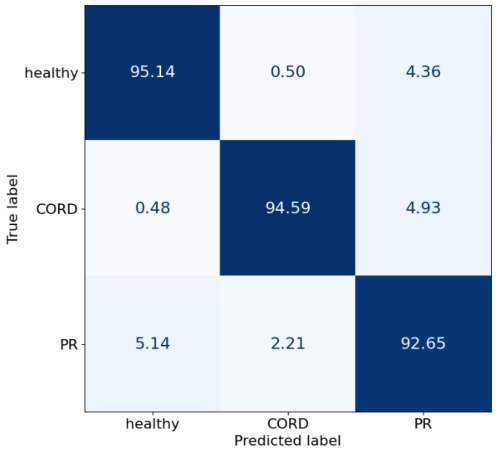


Fig. 7. Confusion matrix for three classes: healthy, CORD and RP cases. It can be seen that the proposed classifier has a little problem in distinguishing between healthy cases and RP, and between RP and CORD. CORD is classified correctly in many cases.

5. DISCUSSION

AI has gained a lot of interest recently, achieving success in various fields of study. In the medical field, many neural network models have been applied for solving challenging tasks concerning retinal diseases [7, 8, 10, 14, 20, 28, 35]. The accuracies depended on various classifiers and datasets consisting of different types of images.

In this study, the Convolutional GRU U-NET model was proposed for the classification of rare eye diseases of rod-dominant (classical RP) and cone-dominant CORD. These two groups were distinguished from healthy cases. A set of experiments was performed, which yielded satisfactory results, exceeding an accuracy of 90%, reaching up to 97.90%. These results of both the proposed model and individual classes confirm that the answer for RQ1 is positive.

The obtained performance results of the proposed the Convolutional GRU U-Net model for RP disorders were compared with the similar outcomes for various eye disease classifications using various networks. In Tab. 7, seven studies were gathered, including RP issues, that used different types of images, datasets and neural approaches. Various types of eye diseases were analysed. The highest accuracy was achieved by MobileNetV2, ImageNet,

InceptionV3 and AlexNet classifiers, while the lowest was achieved by a deep learning approach consisting of DenseNet, Recurrent Neural Network and Fully Connected layer. The accuracy results, obtained in this research, is at the top of the other accuracy values. This indicates that the proposed tool is adequate for the classification of RP disorders. Moreover, based on the state-of-the-art, it can be easily seen that this study made the first attempt to classify RP utilising the proposed neural network.

Analysing the results gathered in Tabs. 2–6, it can be indicated that the proposed classifier can recognise two types of retina dystrophies. The answer for RQ2is definitely affirmative. The proposed Convolutional GRU U-NET model achieved satisfactory results by reaching only 40 epochs, which also gives a positive answer to RQ3. The obtained results clearly indicate that the hypothesis is true and that the proposed Convolutional GRU U-Net can be successfully applied for rare eye diseases.

Tab. 7. Machine learning studies in retinitis pigmentosa (RP) recognition. OCT-optical coherence tomography, CORD-cone-rod dystrophy, UWFP- ultra-widefield pseudocolour, UWFAF-ultra widefield fundus autofluorescence

Classifier	Dataset	Diseases	Accuracy (%)	References
EfficientNet-B7, DenseNet, ResNet-101	574 UWFI images	Diabetic retinopathy, retinal vein occlusion, pathologic myopia, retinal detachment, RP, age-related macular degeneration, vitreous opacity and optic neuropathy	96.44–99.32	[35]
EfficientNet-B7	High-Resolution Fundus	Diabetic retinopathy and RP	85.45	[10]
Own system	Fundus images, OCT	RP	97.90	[7]
Random Forests pixel-based, AdaBoost.M1 pixel-based, Random Forests component-based, AdaBoost.M1 component-based	RIPS, OCT	RP	98.99–99.45	[8]
LCDNet	Fundus camera imaging	Retinal eye diseases	96.52–99.70	[20]
MobileNetV2, ImageNet, InceptionV3 and AlexNet	Kaggle	Glaucoma, maculopathy, pathological myopia and RP	100	[14]
Deep learning approach consists of DenseNet, Recurrent Neural Network and Fully Connected layer	OCT imaging	Neovascular age-related macular degeneration	72	[28]
Convolutional GRU U-Net	UWFP and UWFAF	Healthy, CORD, RP	91.00–97.90	Own

6. CONCLUSIONS AND FUTURE WORKS

Nowadays, the machine learning models have been applied for medical purposes. In this study, a new model was proposed to automatically classify the RP based on Optos images. Three classes were distinguished: healthy, CORD and RP. The Convolutional GRU U-Net was applied for this purpose. The verification of this tool proved that it is a very good solution to classify this rare disease. Each of the mean measures, namely accuracy, precision, sensitivity, specificity and F1 score, produced very high results, above 94%. The learning parameters showed that the dataset was enough to obtain high quality. Additionally, the proposed neural tool was verified using the LOOCV. The obtained results proved the high performance of the developed model. Moreover, according to the authors' knowledge, this is the first type of tool to classify the location of the RP in the retina. The obtained results proved the hypothesis.

The developed Convolutional GRU U-Net shows great potential in practical software implementation to diagnose RP diseases as a supporting tool. It may help to decrease the time of diagnoses. It also may give the recommendations for further treatment. However, this model also has some limitations. It was trained and verified on dataset containing Optors images. Neither temporal nor spatial relationships were taken into consideration. The model requires a high hardware memory capacity. Future studied may extend the classes to perform a more detailed analysis of this rare disease. The proposed neural model may be used for the assessment of the level of the RP. The experiments will be performed focusing on improving the accuracy of the proposed model.

REFERENCES

1. Abeysinghe A, Tohmuang S, Davy J L, Fard M. Data augmentation on convolutional neural networks to classify mechanical noise. Appl. Acoust. 2023:203:109209.
2. Alomar K, Aysel H I, Cai X. Data augmentation in classification and segmentation: A survey and new strategies. J. Imaging. 2023; 9(2): 46.
3. Azad R, Asadi-Aghbolaghi M, Fathy M, Escalera S. Bi-directional ConvLSTM U-Net with densley connected convolutions. 2019. Proc - IEEE/CVF international conference on computer vision workshops.
4. Baratloo A, Hosseini M, Negida A, El Ashal, G. Part 1: Simple definition and calculation of accuracy, sensitivity and specificity. Emergency. 2015;3(2):48-49.
5. Berger W, Kloeckener-Gruissem B, Neidhardt J. The molecular basis of human retinal and vitreoretinal diseases. Prog Retin Eye Res. 2010;29(5):335–75.
6. Bonnici E, Arn P. The impact of Data Augmentation on classification accuracy and training time in Handwritten Character Recognition. Kth Royal Institute of Technology. 2021.
7. Brancati N, Frucci M, Gragnaniello D, Riccio D, Di Iorio V, Di Perna L. Automatic segmentation of pigment deposits in retinal fundus images of Retinitis Pigmentosa. Comput. Med. Imag. Graph. 2018;66:73-81.

8. Brancati N, Frucci M, Gragnaniello D, Riccio D, Di Iorio V, Di Perna L, Simonelli F. Learning-based approach to segment pigment signs in fundus images for retinitis pigmentosa analysis. Neurocomputing. 2018,308:159-171.
9. Chen JX, Jiang DM, Zhang YN, A hierarchical bidirectional GRU model with attention for EEG-based emotion classification, IEEE Access. 2019;7:118530-118540.
10. Das H, Saha A, Deb S. An expert system to distinguish a defective eye from a normal eye. Proc - 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). IEEE. 2014:155-158.
11. Fahim AT, Daiger SP, Weleber RG. Nonsyndromic retinitis pigmentosa overview. 2017: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, Amemiya A, eds.Gene Reviews. Seattle: University of Washington.
12. Gill JS, Georgiou M, Kalitzeos A, Moore AT, Michaelides M. Progressive cone and cone-rod dystrophies: Clinical features, molecular genetics and prospects for therapy. Br. J. Ophthalmol. 2019;103(5): 711-720.
13. Graves A, Mohamed AR, Hinton G. Speech recognition with deep recurrent neural networks. Proc - IEEE Int. Conf. Acoust., Speech Signal Process. 2013;38:6645–6649.
14. Guo C, Yu M, Li J. Prediction of different eye diseases based on fundus photography via deep transfer learning. J. Clin. Med. 2021;10(23):5481.
15. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proc - IEEE conference on computer vision and pattern recognition. 2017:4700-4708.
16. Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. Lancet. 200618;368(9549):1795-809. doi: 10.1016/S0140-6736(06)69740-7
17. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proc - International conference on machine learning. 2015:448-456.
18. Jain L, Murthy H S, Patel C, Bansal D. Retinal eye disease detection using deep learning. Proc - Fourteenth International Conference on Information Processing (ICINPRO). IEEE. 2018:1-6.
19. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.
20. Liu T Y A, Ling C, Hahn L, Jones C K, Boon C J, Singh M S. Prediction of visual impairment in retinitis pigmentosa using deep learning and multimodal fundus images. Br. J. Ophthalmol. 2022.
21. Masumot H, Tabuchi H, Nakakura S, Ohsugi H, Enno H, Ishitobi, N., Ohsugi E, Mitamura Y. Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultrawide-field images. PeerJ. 2019;7:e6900.
22. Merin S, Auerbach E. Retinitis pigmentosa. Surv. of Ophthalmol. 1976; 20(5):303-46. doi: 10.1016/s0039-6257(96)90001-6
23. Monaghan T F, Rahman S N, Agudelo C W, Wein A J, Lazar J M, Everaert K, Dmochowski R R. Foundational statistical principles in medical research: sensitivity, specificity, positive predictive value, and negative predictive value. Medicina. 2021;57(5):503.
24. Oishi A, Miyata M, Numa S, Otsuka Y, Oishi M, Tsujikawa A. Widefield fundus autofluorescence imaging in patients with hereditary retinal degeneration: a literature review. Int. J. of Retina Vitr. 2019;12(5) (Suppl 1):23. https://doi.org/10.1186/s40942-019-0173-z.
25. Piri N, Grodsky JD, Kaplan HJ. Gene therapy for retinitis pigmentosa. Taiwan J. Ophthalmol. 2021;11(4):348-351.
26. RetNet Retinal Information Network. https://sph.uth.edu/retnet/ [6.06.2023]
27. Robson AG, Egan CA, Luong VA, Bird AC, Holder GE, Fitzke FW. Comparison of FAF with photopic and scotopic fine-matrix mapping in patients with retinitis pigmentosa and normal visual acuity. Invest. Ophthalmol. Vis. Sci. 2004;45(11):4119-4125.
28. Romo-Bucheli D, Erfurth U S, Bogunović, H. End-to-end deep learning model for predicting treatment requirements in neovascular AMD from longitudinal retinal OCT imaging. IEEE J. Biomed. Health Inform. 2020;24(12):3456-3465.
29. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Proc - 18th International Conference, Munich, Germany. October 5-9. Proceedings. Part III. 2015; 18:234-241. Springer International Publishing.
30. Schmitz-Valckenberg S, Holz FG, Bird AC, Spaide RF. Fundus autofluorescence imaging: review and perspectives. Retina. 2008;28(3):385-409.
31. Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Adv. Neural. Inf. Process. Syst. 2015;28.
32. Shorten C, Khoshgoftaar T M. A survey on image data augmentation for deep learning. J. Big Data. 2019;6(1):1-48.
33. Skublewska-Paszkowska M, Powroznik P. Temporal Pattern Attention for Multivariate Time Series of Tennis Strokes Classification. Sensors. 2023;23(5):2422.
34. Song H, Wang W, Zhao S, Shen J, Lam KM. Pyramid dilated deeper convlstm for video salient object detection. Proc - European conference on computer vision (ECCV). 2018:715-731.
35. Sun G, Wang X. Xu L. Li C. Wang W. Yi, Z., Luo H, Su Y, Zheng J, Li Z, Chen Z, Zheng H, Chen, C. Deep learning for the detection of multiple fundus diseases using ultra-widefield images. Ophthalmol. Ther. 2023;12(2):895-907. https://doi.org/10.1007/s40123-022-00627-3
36. Tee JJ, Smith AJ, Hardcastle AJ, Michaelides M. RPGR-associated retinopathy: clinical features, molecular genetics, animal models and therapeutic options. Br J Ophthalmolol. 2016;100(8):1022-7. doi: 10.1136/bjophthalmol-2015-307698
37. Wong S C, Gatt A, Stamatescu V, McDonnell M D. Understanding data augmentation for classification: when to warp? Proc - International conference on digital image computing: techniques and applications. IEEE. 2016:1-6.
38. Yang S, Xiao W, Zhang M, Guo S, Zhao J, Shen F. Image data augmentation for deep learning: A survey. 2022. arXiv preprint arXiv:2204.08610.

"Approval of the Ethic Committee of the Medical University of Lublin has been obtained (nr KE-0254/260/12/2022). All procedures conformed to the tenets of the Declaration of Helsinki."This project has been done within the "Staż za miedzą" restricted grant funded by the Medical University in Lublin, Poland (Chair and Department of General and Pediatric Ophthalmology, Medical University of Lublin, Poland and Faculty of Electrical Engineering and Computer Science, Department of Computer Science, Lublin University of Technology, Lublin, Poland).The study was carried out as a part of the project "Lubelska Unia Cyfrowa – Wykorzystanie rozwiązań cyfrowych i sztucznej inteligencji w medycynie – projekt badawczy", no. MEiN/2023/DPI/2194.

Maria Skublewska-Paszkowska: https://orcid.org/0000-0002-0760-7126

Pawel Powroznik: https://orcid.org/0000-0002-5705-4785

Robert Rejdak: https://orcid.org/0000-0003-3321-2723

Katarzyna Nowomiejska: https://orcid.org/0000-0002-5805-8761