UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
UNIDADE ACADÊMICA DE SISTEMAS E COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

MARCUS ANTONIO ROCHA TENORIO

PROPOSAL OF A LOW-COST DEVICE TO SUPPORT REMOTE DIABETIC RETINOPATHY DETECTING BASED ON FUNDUS IMAGES

CAMPINA GRANDE - PB

2020

MARCUS ANTONIO ROCHA TENORIO

PROPOSAL OF A LOW-COST DEVICE TO SUPPORT REMOTE DIABETIC RETINOPATHY DETECTING BASED ON FUNDUS IMAGES

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Professor Dr. Herman Martins Gomes.

CAMPINA GRANDE - PB

2020

T312p Tenorio, Marcus Antonio Rocha.

Proposal of a low-cost device to support remote diabetic retinopathy detecting based on fundus images. / Marcus Antonio Rocha Tenorio. – 2020.

12 f.

Orientador: Prof. Dr. Herman Martins Gomes.

Trabalho de Conclusão de Curso - Artigo (Curso de Bacharelado em Ciência da Computação) - Universidade Federal de Campina Grande; Centro de Engenharia Elétrica e Informática.

1. Retinopatia diabética. 2. Dispositivo de baixo custo. 3. Tecnologia aplicada à saúde. 4. Oftalmologia - tecnologia. 5. Fundo de olho - imagens. 6. Detecção remota de retinopatia diabética. 7. Imagens de fundo de olho. 8. Redes neurais artificiais. I. Gomes, Herman Martins. II. Título.

CDU:004(045)

Elaboração da Ficha Catalográfica:

Johnny Rodrigues Barbosa
Bibliotecário-Documentalista
CRB-15/626

MARCUS ANTONIO ROCHA TENORIO

PROPOSAL OF A LOW-COST DEVICE TO SUPPORT REMOTE DIABETIC RETINOPATHY DETECTING BASED ON FUNDUS IMAGES

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

BANCA EXAMINADORA:

Professor Dr. Herman Martins Gomes
Orientador – UASC/CEEI/UFCG

Professor Dr. Fábio Jorge Almeida Morais
Examinador – UASC/CEEI/UFCG

Professor Dr. Tiago Lima Massoni
Disciplina TCC – UASC/CEEI/UFCG

Trabalho aprovado em: 2020.

CAMPINA GRANDE - PB

Proposal of a low-cost device to support remote diabetic retinopathy detecting based on fundus images

Marcus A.R Tenorio
marcus.tenorio@ccc.ufcg.edu.br
Federal University of Campina Grande
Campina Grande, Paraiba, Brazil

Herman Martins Gomes
hmg@computacao.ufcg.edu.br
Federal University of Campina Grande
Campina Grande, Paraiba, Brazil

ABSTRACT

Diabetes causes several problems, including diabetic retinopathy, which when discovered belatedly can lead to total blindness. Brazil is also the 8th largest country in the world, with conurbation problems and an increase in diabetes diagnosis in the past 10 years. In this context, the present work aims to propose a low-cost prototype to support the diagnosis of diabetic retinopathy based on *fundus* examinations images so that physicians are able to perform early diagnosis in remote locations.This prototype should allow for early detection and treatment *in loco*, thus increasing the chances of a positive outcome for the patients. First we studied technical aspects relevant to the proposal such as physiological aspects of diabetic retinopathy, Artificial Neural Networks, Accelerated and Edge computing. Our methodology consisted in a comparison of embedded hardware with capabilities to perform complex computations, a survey of models for the classification of diabetic retinopathy and available databases, including research choices. Artificial Neural Networks to identify diabetic retinopathy were evaluated in our low-cost embedded system in terms of accuracy. The accuracy must be enough to determine the priority of the patient's case for treatment. This work reached accuracy levels around 84% with a low cost system and less computational power, positioning itself well in the state of the art of systems within greater computational power. The results indicate that the platform is indeed low-cost and suitable for this application.

ACM Reference Format:
Marcus A.R Tenorio and Herman Martins Gomes. 2020. Proposal of a low-cost device to support remote diabetic retinopathy detecting based on fundus images . In . , 9 pages.

1 INTRODUCTION

According to a report by the Surveillance of Risk and Protection Factors for Chronic Diseases by Telephone Survey (VIGITEL) [3] from the Ministry of Health, the number of Brazilians with diabetes between the years 2006 and 2016, increased by 61.8%. The disease now affects a total of 8.9% of the population against 5.5% in a previous survey [2]. This index is more accentuated in females, where the disease reaches 9.9% compared to 7.8% of the male population. According to the VIGITEL survey [3], this trend of growth of diabetes is observed worldwide, that occurs due to the influence of factors such as aging of the population, changes in eating habits and lack of physical activity.

Among the different complications of diabetes there is the Diabetic Retinopathy (DR), which is the main cause of blindness in the world's working-age population [40]. DR is a disease that affects the small vessels of the retina, the region of the eye responsible for the formation of images sent to the brain. The onset of diabetic retinopathy is mainly related to the duration of diabetes and uncontrolled blood glucose [22]. When diabetes is not controlled, hyperglycemia triggers several changes in the body that, among other damages, leads to dysfunction of the retinal vessels [5].

Typically, diabetic retinopathy can be detected from [6]:

- Examination of visual acuity: assesses the ability to distinguish small letters from a 20 feet distance;
- Fundoscopy: important to assess the stage of retinopathy;
- Angiography: reveals neovascularization, aneurysms, and hemorrhage;
- Optical tomography: reveals macular edema and macular degeneration.

All the above mentioned examinations require specialized (and high cost) equipment that are commonly found in large city centers [26], but not in remote locations. Brazil is the sixth more populous country in the world [1], but the population density is only $23,8\ inhab./km^2$. The "Sistema Único de Saúde" (SUS), Brazil's integrated health system, planned to decentralize the primary care at the municipal level, wasn't implemented as planned. This means that the primary care is not very accessible and specialized treatment is often only available in city centers [25].Cunha et al. [7] noted that the geographical accessibility to health care services still lacks improvement. This leads to the lack of access to medical resources in the most remote areas.

Despite the growth of the medical informatics field, the implementation of Telemedicine in Brazil is still a big challenge [21]. Telemedicine has an interdisciplinary nature, that is, its development requires the participation of several areas of knowledge – medical, ICT, microelectronics, information technology, telecommunications, equipment, among others –, which reinforces the need for a systemic perspective and actions coordinated among different decision-making bodies, with the participation of industry, academia, scientific and technological institutions, class associations, among other relevant agents in the innovation process. Thus, our long term goal is to fill that gap by

, ,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00

[1] https://pt.wikipedia.org/wiki/Lista_de_pa%C3%ADses_por_popula%C3%A7%C3%A3o

bringing industry, academia and scientific institutions all together in order to better approach the DR in the current Brazilian scenario.

This work proposes to create a low cost platform to be used in health units in Brazil where access to more advanced systems is not so easy, in addition to helping technicians or administrative assistants with minimal training to be able to create a queue of priority using images from an eye exam so that they can optimize the service queue of these people in the most critical situation for specialized care.

Systems for detecting diabetic retinopathy almost always depend on a specialist who may be wrong, as the quality of the diagnosis is intrinsically linked to the cost of the device.

However, the simplest device for detecting diabetic retinopathy is around hundreds of Reais, it needs a specialist for its use.

Which makes it ineffective in most environments where we are proposing to use our solution.

Whereas more sophisticated devices can reach tens of thousands of Reais, as they almost always need a proprietary and high-cost system for all their functioning.

In this work, we evaluated two candidate hardware platforms to run the diabetic retinopathy classifier: a Jetson Nano[2] or a Raspberry Pi 4[3] A comparative study was carried out in order to choose the best suited platform.

Several image datasets were used to evaluate existing diabetic retinopathy classifiers: A dataset of images obtained by an ophthalmologist from Campina Grande (Brazil) [31], which were anonymized; as well as the KAGGLE [15] image base, messidor dataset and the Indian base. These datasets are described in the methodology section.

The paper is structured as follows: in Section 2 the background of the research is presented. Next, in Section 3 the methodology regarding experiments is presented. After that, in Section 4 we present the setup of our experiments as well as the requirements to our results which are presented in Section 5, for the different scenarios. Finally, Section 6 contains the conclusions and proposals of future work.

2 BACKGROUND

Given the multidisciplinary nature of the topic, the follow subsections will give more specific details on the different concepts of the research. First, we give an introduction on DR and how it is approached currently. Next, we present an overview on Artificial Neural Networks and some state of the art models to address the DR problem. Following, we discuss embedded systems solutions currently available and the reason of choosing the Jetson Nano board. Finally, a review of related work is presented.

2.1 Diabetic Retinopathy

Diabetic retinopathy is the most frequent cause of new cases of blindness among adults aged 20–74 years [9]. Within the first two decades of the disease, nearly all patients with type 1 diabetes and >60% of patients with type 2 diabetes have retinopathy. The retina, depicted in Figure 1, is the light-sensitive layer of cells at the back of the eye that converts light into electrical signals. The signals are sent to the brain which turns them into the images one sees.

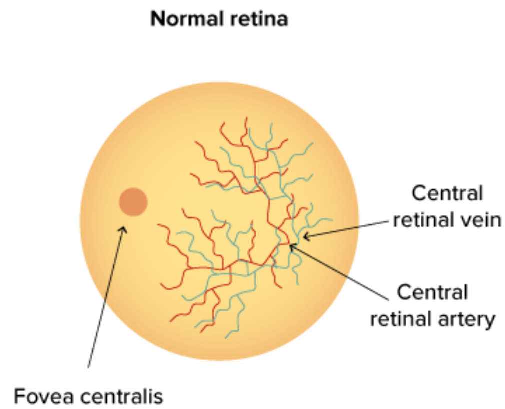

Figure 1: Normal retina, adapted from [4]

The retina needs a constant supply of blood, which it receives through a network of tiny blood vessels [30], as depicted in Figure 2. Over time, a persistently high blood sugar level can damage these blood vessels in 3 main stages [30]:

- background retinopathy – when some lumps appear in the blood vessels, which may bleed but don't usually affect the vision;
- pre-proliferative retinopathy – when the blood vessels are more affected, including significant increase bleeding into the eye;
- proliferative retinopathy – when scar tissue and new blood vessels develop on the retina causing some loss of vision.

Therefore, diabetic retinopathy is a pathology consequent from the persistent high levels of blood sugar caused by diabetes. However, if picked up early, lifestyle changes and/or treatment can stop it from getting worse [1].

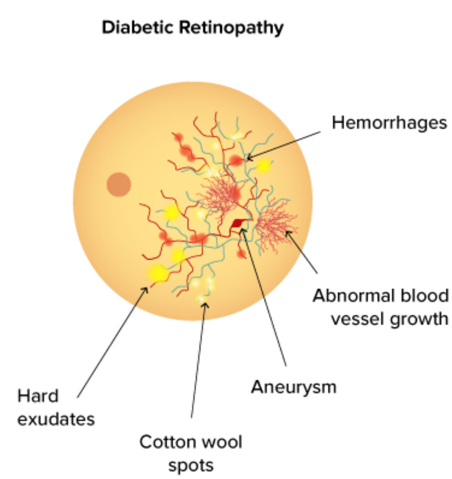

Figure 2: Diabetic retina, adapted from [4]

[2]https://developer.nvidia.com/embedded/jetson-nano-developer-kit
[3]https://www.raspberrypi.org/products/raspberry-pi-4-model-b/

DR is present in both diabetes types. Type 1 diabetic patients do not usually develop vision-threatening retinopathy in the first 3–5 years of diabetes or before puberty. However, during the next two decades, nearly all type 1 diabetic patients develop retinopathy [11]. For the type 2 diabetes patients, up to 21% have retinopathy at the time of first diagnosis of diabetes, and most develop some degree of retinopathy over time [11].

2.2 Artificial Neural Networks

According to Haykin [12] a neural network can be defined as:

> "...a massively parallel distributed processor made up of simple processing units that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:
> 1. Knowledge is acquired by the network from its environment through a learning process.
> 2. Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge. "

The concept of ANNs is based on a collection of connected nodes called artificial neurons and their connections, called edges. The edges and neurons have weights, a constant value that will adjust as learning proceeds. The weight increases or decreases the strength of the connection between nodes.

2.2.1 Deep Neural Networks. ANNs when composed of many layers, are called deep neural networks(DNN). Due to the depth of these networks it is possible to learn a sequence of functions that perform the transformation of vectors. [37]

2.2.2 Convolutional Neural Network. In deep learning, a convolutional neural network (CNN) is a class of deep neural networks, most commonly applied to analyzing visual imagery [38]. CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage.

2.3 Accelerated Computing for AI

Several studies of deep learning applications, were found more successful when in use of some sort of hardware acceleration [10]. Developments of neural networks and algorithms for modern computer vision have been observed as computational power is getting accessible. On the other hand, data-driven and cloud computing based systems face some limitations at edge devices in real-world scenarios [24].

Hardware acceleration is the use of specific hardware to perform certain tasks more efficiently or faster than with general purpose processors.

Traditionally, processors are sequential, that is, instructions are executed one by one. Most algorithms are designed to work sequentially. Hardware acceleration aims to improve the execution of specific algorithms allowing the use of parallelism.

The hardware that performs the acceleration, when it is separated from the CPU, is called a hardware accelerator, or more specifically, for example, cryptographic accelerator, floating point accelerator, graphics accelerator, etc.

There are several use cases of acceleration in hardware such as calculations of meteorological problems, operations with large numbers, calculations for bioinformatics, among others.

It should be noted that there is a trade-off between efficiency and flexibility: the more efficient the process, the less flexible the change is.

Some os the most well-known hardware accelerator are FPGAs and GPUs. We'll present them in the next sections.

2.3.1 FPGA. An FPGA (Field Programmable Gate Array) [23] is an integrated circuit that contains a large number of identical logical units. In a simple way, an FPGA is similar to a digital prototyping board, that is, we have a large amount of components (logic gates) that can be assembled in the most diverse ways for the most diverse purposes. A FPGA usually has three main components:

(1) Look-up Tables (LUT)
(2) Flip-Flops
(3) Rotation matrix

They work in a coordinated manner to create a flexible device, suitable for diverse applications, that by definition will be programmable.

This programming is done in different ways, state machines, schematic and the most used: hardware description languages (HDL).

All of this makes it easier to design a digital circuit, without the welding and other factors much more critical to the creation of an integrated circuit.

For those reasons, FPGA is currently used as a part of the integrated circuit development flow for creating an ASIC (Application Specific Integrated Circuit), since the cost of an FPGA is much lower than the construction of a specific ASIC for any application.

It is important to note that, unlike a programming language like C, JAVA, Python and others, HDL depends on the FPGA being used and needs extra electronics knowledge to understand and execute the code.

2.3.2 GPU. GPUs (Graphics Processing Unit) [17] are graphics processing units, and this term was disclosed by NVDIA at the launch of its GFORCE board in 1999. The GPU market came up with because of the high demand for games that needed specific, and in most cases parallel, processing.

This trend is also in line with the recent increase in so-called multicore architectures, processors that have more than one processing core, even if the clock rate is lower than the old processors. This approach is called parallelization, where several cores are used to perform tasks.

Usually GPUs have a much larger number of processors that facilitate the parallelization of processes. Although the names multicore and manycore are widely used to label the CPU and GPU architectures, respectively, it is necessary to consider that the processing cores of each platform have different characteristics.

A CPU core is typically more robust and designed for more general tasks and more complex control flows, focusing on the quick execution of sequential tasks. A GPU core has relatively less sequential computational capacity, but being optimized for parallelism in data processing.

Thus, it presents a more elementary logic control scheme, whose purpose is to favor the flow of the parallel application.

2.4 Edge computing

In the late 1990s, Akamai introduced content delivery networks (CDNs) to accelerate web performance [8]. A CDN uses nodes at the edge, close to users, to: (*i*) prefetch and cache web content, (*ii*) perform some content customization, such as adding location-relevant advertising [34].The concept of CDN mixes a lot with that of edge computing. We can define edge computing when there is a device that captures and processes data from sensors, actuators and other data sources close to where they occur. Cloud computing than was the big player responsible for allowing Akami's work to be extended and now globally implemented also in small devices.

The advances in microelectronics as well as in computing performance allowed the appearance of low-cost systems such as Arduino[4], Raspberry Pi[5] and, most recently, the Jeston Nano development board by NVIDIA, presented in figure 3. These development boards have embedded microprocessors that allows computing at portable devices. Prices range usually from couple of dozens to few hundred dollars.

Figure 3: Jetson Nano by NVIDIA

Among the many available options, the Jetson Nano was chosen as the base system for the experiments related in this paper. It adds low-cost high performance (with a GPU) in a credit card sized development board. Ideal for a portable device that requires high computational capabilities at a low-cost for remote purposes.

2.5 Related Work

Back in 1992, *Lairson et al.* [19] conclude their work by stating that primary-care screening with retinal photographs through pharmacologically dilated pupils for diabetic retinopathy is an appropriate and cost-effective alternative to screening by an ophthalmologist in this setting [19]. This originated a trend on analysing retinal photographs that persists up until today.

In a recent study [26], the authors suggested that the ideal screening technology for DR must be portable, noninvasive, reliable, and easy to use by relatively unskilled persons. Testing must be deployed in areas with sufficient volume of patients that resources spent on travel cover the cost reduction in preventing blinding disease. The objective of screening programs is to identify individuals who will benefit from sight saving laser therapy. Also, *Leese et al.* [20] identified that the mobile diabetic eye screening program detected a greater prevalence of advanced retinopathy in diabetic patients living in rural areas. Patients in rural areas were also more likely to need urgent laser photocoagulation [20].

With that in mind, solutions like the Horus DEC 200[6] were developed. *Quellec et al.* [32] assessed the suitability of a low-cost, handheld, nonmydriatic retinograph, namely the Horus DEC 200, for diabetic retinopathy (DR) diagnosis – where two factors were considered: ease of image acquisition and image quality. It was observed that the Horus can be used to screen DR, but at the cost of longer examination times and higher proportions of patients referred to an ophthalmologist due to inadequate image quality.

In 2013, *Prasanna et al.* proposed a handheld system [28] based on an Android mobile smartphone. It is based on Support Vector Machines (SVM) and Logistic Regression but due to its limited computing capacity and lack of GPU, other classifiers or techniques may not be implemented.

Another mobile-based solution was proposed in 2018 [33] assessing the Artificial Intelligence DR screening software (EyeArtTM)[7] – a SaaS solution designed to identify DR, referable DR (moderate non-proliferative DR or worse and/or DME) or STDR. Despite the satisfactory results, the algorithms and solutions implemented are proprietary, thus enhancing costs specially to countries where the currency is depreciated with respect to the American dollar.

3 METHODOLOGY

The methodology adopted in order to create and validate the proposed prototype was: (*i*) survey of the existing edge computing boards on the market, (*ii*) comparison between these boards and chose the one that could best suit the application context, (*iii*) implementation of different DR identification networks in the chosen board, (*iv*) performance comparison of the networks in terms of accuracy and processing time, (*v*) using private and public datasets.

The following sections discuss the methodology steps presented above.

3.1 Survey of boards

During the exploratory stage, several boards were taken into account for this research, such as Intel's DE0-nano[8] and other similar FPGAs due to their accelerated computing capacity and ease of design for digital systems.

Other available solutions are Visual Processing Units (VPU) such as Intel's Neural Stick, that can be connected to a USB port in order to provide accelerated computing.

[4]https://www.arduino.cc/

[5]https://www.raspberrypi.org/

[6]http://www.miis.com.tw/product01.php?no=84

[7]https://www.eyenuk.com/en/products/eyeart/

[8]https://www.terasic.com.tw/cgi-bin/page/archive.pl?No= 593

There were also solutions from NVIDIA, with GPUs, such as the Jetson Nano and the Jetson TX2 boards. They differ from the previous two options because there's no hardware description language involved and doesn't rely on another system to properly function (as the stick). They're complete embedded solutions with WiFi capabilities and allows the user to chose which OS to use.

Following this, there are the Raspberry Pi family. They're complete development boards, as the NVIDIA's, but at a much lower cost. The most recent model, Raspberry Pi 4, if the first to have a GPU. This make it the only suitable candidate from Raspberry family.

Since the objective is to replicate pre-existing models for the detection of diabetic retinopathy, the adaptation of state-of-the art deep learning algorithsm to new architectures / programming languages such as HDL was not plausible. Also, the use of an external component such as an Intel Neural Stick would increase the complexity of the system. Therefore, the three main candidates are: Jetson TX2, Jetson Nano and Raspberry Pi 4 (RPI 4).

3.2 Comparing boards

After selecting the board candidates, the following comparative table was obtained from their specifications.

	Jetson TX2	**Jetson Nano**	**RPI 4**
RAM	8 GB (128 bit)	4 GB	4 GB
CPU	Denver 2 and Quad-Core ARM Cortex-A57	Quad-Core ARM Cortex-A57 @ 1.42 Ghz	Quad-core ARM Cortex-A72 @ 1.5 Ghz
GPU	NVIDIA Pascal	NVIDIA Maxwell	Broadcom VideoCore VI (32-bit)
Cost (USD)	299	99	55

After careful consideration of the information presented on the table above, it was noticed that the following points were crucial to this choice:

(1) **Cost of the platform** - As we are talking about a low cost platform that must be used in different health units, we need a setup that complies to the budget of low-budget cities. Very often, the best solution is not always the cheapest.
(2) **Availability** - The board must be easy to obtain. International shipping and custom duties are usually limitations for many health devices.
(3) **Computational power** - Since the solution relies on Deep Learning models, the computation power must be enough to compute those models locally in a reasonable time.
(4) **Replicability of known models** - The board must also have and environment where most modern programming languages used for deep learning are supported and somewhat easy to use.

Therefore, based on the above mentioned criteria, we chose to move forward with the Jetson Nano board.

3.3 DR identification networks

An experiment model was built thinking about the reproduction of already known models. In this scenario, the images should be received by the system, anonymized and pre-processed. The algorithms then will have the pre-processed data to work with and provided the output.

After a state of the art review in the literature [39] the following networks were chosen for comparison.

3.3.1 AlexNet. AlexNet is a convolutional neural network created in 2012 [18]. When the domain of pattern and image processing was introduced in 2012, it still worked with "small" datasets, but with the growth of computational power and the need for more accurate results, the need to create a network that could process a greater amount of data. images like those made available on ImageNet, emerged.

Its standard and non-optimized architecture is made up of five layers and in each one a rectifier function is used to activate neurons, after this layer a pooling layer is found.

The AlexNet network adopts the ReLU activation function, to perform training several times faster than other functions.

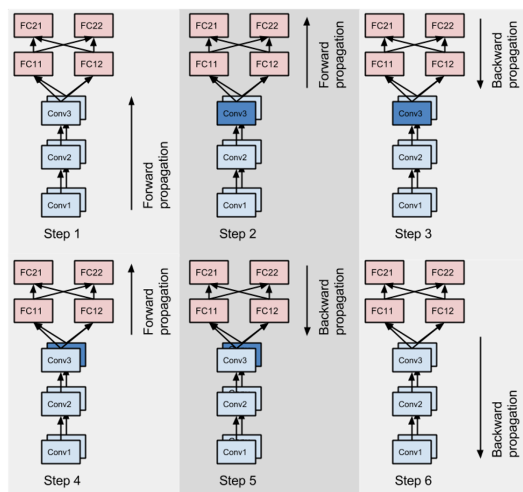

Figure 4: AlexNet Architecture

3.3.2 VGG16. VGG16 (Visual Geometric Group - 16) was proposed in 2014 by Karen Simonyan and Andrew Zisserman [35] is a network that has sixteen layers. Some authors [39] consider it an evolution of a Convolutional neural network. The great aspect of VGG16 is the simple way of building its hidden layers.

VGG is know to explore filters with 3x3 dimensions with small receptive fields for a deep architecture, and show that 7x7 filters, that were used in Alexnet, can be replaced by a sequence of 3 3x3 filters, reducing the number of parameters.

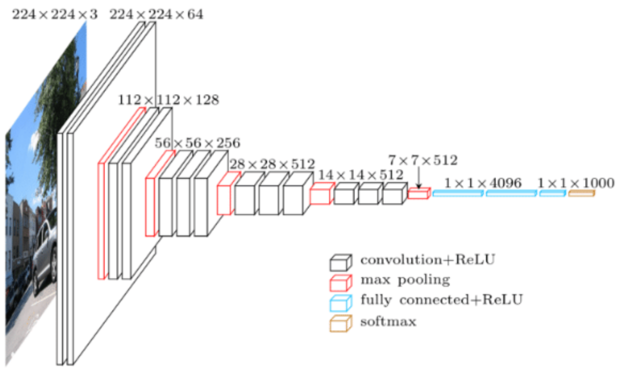

Figure 5: VGG Architecture

3.3.3 ResNet50. ResNet [13] was created in 2016. This model is a convolutional neural network made up of more than 150 layers. Its learning process follows that its deep neural networks when reshaping the layers with residual functions such as reference to layer entries. Another important feature of ResNet is that it has an intense use of BatchNormalization layers and does not have dense layers except its output layer. The architecture is depicted in Figure 6.

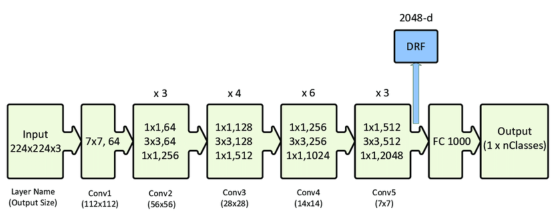

Figure 6: ResNet50 Architecture

The authors reformulated the layers as learning residual functions, with reference to the layer inputs, instead of learning unreferenced functions [14]. They also provided empirical evidence that these residual networks are easier to optimize, and may improve accuracy. Using the ImageNet dataset, the authors evaluated residual nets with a depth of up to 152 layers – 8x deeper than VGG nets but still having lower complexity.

3.3.4 InceptionNet. Inception [36] is a convolutional neural network that aims to decrease the number of parameters with its Inception block, without losing its efficiency. This network stands out for having departed from the known standard of sequential neural networks presenting forks within the network, and the use of more than one classifier. T

The Inception network, as depicted in Figure 7, is composed of 9 blocks called “inceptions blocks” that propagate the information of a layer in 4 data streams, in which each stream is responsible for extracting information at different levels of abstraction.

3.4 Datasets

For validation of our model, two different datasets were used which are described in the subsections below.

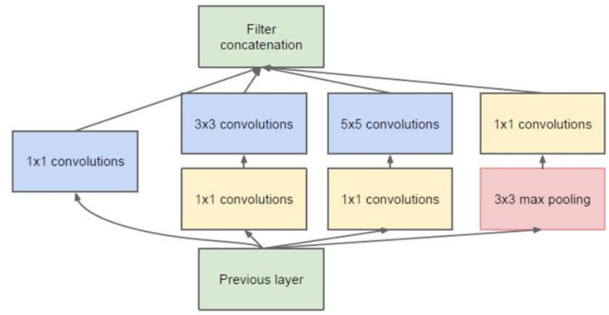

Figure 7: InceptionNet Architecture

Stages are the different levels of diabetic retinopathy that can occur in a human eye, with stage 0 being a healthy eye and stage 5 being an eye in an advanced stage of diabetic retinopathy.

3.4.1 Kaggle. [15]

Within the means of researching medical images and especially diabetic retinopathy, the dataset constituted for kaggle in your competition is one of the best known and used for training networks for the detection of diabetic retinopathy, the base has a total of 53,576 images for testing, 35126 images are in their training set, these are properly marked with the following classification:

Stage	Diabetic Retinopathy Classification	Images
0	No Diabetic Retinopathy	25.810
1	Mild Severe	2443
2	Moderate Severe	5292
3	Severe	873
4	Proliferative Diabetic Retinopathy	708

Table 1: Kaggle dataset description.

3.4.2 Campina Grande, Brazil. This was a private dataset provided by a Dr. Fabio Queiroz from Campina Grande, even though it has only 135 images that are labeled in the different stages described in the previous subsection, this dataset is particularly interesting because it is an original base and has been labeled by a local doctor.

Stage	Diabetic Retinopathy Classification	Images
0	No Diabetic Retinopathy	62
1	Mild Severe	25
2	Moderate Severe	12
3	Severe	26
4	Proliferative Diabetic Retinopathy	11

Table 2: Dr.Fabio Queiroz dataset description.

4 EXPERIMENTS SETUP

The main objective of the experiments was to check if the Jetson Nano would be able to run models for detecting diabetic retinopathy and obtain significant results for be quoted as a low-cost alternative to our base case.

The following setup was used: Jetson Nano 4GB with Linux4Tegra.

The validation of the networks was carried out at Jetson Nano using widely used frameworks from the community some of then with native implementations of the above-cited networks such as keras, pytorch and tensorflow.

The model used for AlexNet[9] and InceptionNet[10] was the implementation made in the PyTorch library.

For the VGG16[11] and Resnet50[12] implementation, we use the model available in the Keras library

The images were pre-processed locally and then evaluated by the Alexnet, VGG16, ResNet50 and InceptionNet networks implemented at the Jetson Nano board, as depicted in Figure 8.

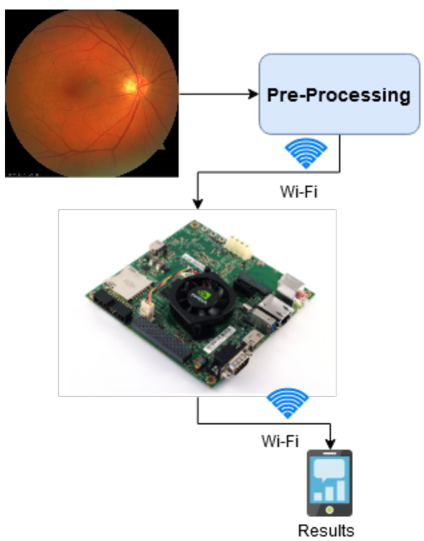

Figure 8: Experiment Setup

The output of the networks may then be made available for any consumer on the internet through a simple REST API.

4.1 Preprocessing

We had to preprocess the images from both the kaggle datase and from our Campina Grande dataset in order to obtain a normalization of the images for training our networks.

The preprocessing has become necessary since the images are of patients from the most different stages of diabetic retinopathy.

Therefore, a preprocessing using the OpenCV library was necessary to make the differences in the images that showed signs of diabetic retinopathy more expressive.

These images were resized to 128x128, to be processed efficiently by Jetson Nano.

In both datasets we have images with a high resolution.

First, images were cropped to remove the borders.

[9]https://pytorch.org/hub/pytorch$_vision_alexnet/$

[10]https://pytorch.org/hub/pytorch$_vision_inception_v3/$

[11]https://keras.io/api/applications/vgg/

[12]https://keras.io/api/applications/resnet/

After that, we studied the cases known in the literature for preprocessing the images. [29]

And especially in our case, where detecting micro-aneurysms is one of the most important features.

The green channel was removed from the images already treated and with a lower resolution to increase the visibility of micro-aneurysms.

4.2 Overfitting

We know about the existing problem of overfitting in neural networks, several works on diabetic retinopathy point out this problem since most images always belong to the so called stage 0 (no DR) of our classification, for this weights had to be adapted over time.

One of the most common and also used ways within the field of diabetic retinopathy research is the use of Data Augmentation Techniques.This was done in our work, using rotations in the images after preprocessing.

5 EXPERIMENTAL RESULTS

The results expressed in the table 3 below show the accuracy of the model after training, the KAGGLE dataset training base were first used [15].

The datasets utilized to give the reality of how the model is performing was the datasets described in Tables 1 and 2 of this document

Stage	AlexNet	VGG16	ResNet50	InceptionNet
0	87.6%	87.3%	86.9%	85.2%
1	86.2%	85.6%	80.1%	84.6%
2	84.1%	86.9%	84.7%	83.5%
3	81.2%	79.6%	81.7%	82.5%
4	83.8%	80.1%	79.2%	84.9%

Table 3: Experimental Results.

These results compared within the literature as [16], provide one of the necessary validations for this work, as we reach a high level of accuracy for the standards of a low cost system compared too a high-end experiment such as the one cited in [16].

In this work, a better accuracy was found for networks similar to those used in this work, where the average accuracy of these networks was around 97 %, while the average of our networks was around 85 %.

However, the experimental setup of the aforementioned article is much more expensive than ours and served as a floor reference for this article.

As the focus of this work was efficiency and simplicity of the proposed solution, the results shows that as an entry system for prioritizing patients without the assistance of a specialist, the answer is quite satisfactory.

As we know and mentioned earlier, the Medical Doctor/population ratio in our country is still not ideal.

So, even without the presence of a specialist, the platform may help even as a second opinion to the specialist or as a prioritization tool.

6 CONCLUSION AND FUTURE WORK

In this work, we investigate the possibility of creating a low-cost system for detecting diabetic retinopathy, even if solutions such as Phelcom Eyer[13], already available on the market.

That work represents a relevant step towards a the creation of an environment to support and detect early cases of diabetic retinopathy.

Its initial value of U$ 5000 makes it very affordable compared to the usual office solutions that cost U$ 20000.

However, the objective of this proposal is to create a facilitator between the most distant locations and its prioritization in attendance to a larger health center, that said, the proposal is viable due to the low cost of the U$ 99 board and its ease of implementation, in addition to portability together with a d-eye camera that costs approximately U$ 435[14], the cost of our system is still well below the solutions available in the market today.

A future work is the creation of educational materials to use the solution, in addition to the creation of a web support system to facilitate the exchange of knowledge between specialists and lay users.

A threat to the validity of the experiments is their relatively high variability, a future work would be the execution of new experiments in comparison with other networks and a better statistical analysis.

Use of new validations such as accuracy, awareness and effectiveness for example using Quadratic Kappa Metric [15] is a relevant future work.

Unlike using the only accuracy, this metric adapts to the issue of false positives and false negatives.

When we are talking about systems used for medical assistance, a false negative or a false positive can generate a great deal of human damage, so this metric helps to understand this relationship to which the accuracy is only a result of successes and errors.

The result of the work is a low-cost system that can be easily replicated in distant cities, thus fulfilling its social objective of bringing healthcare closer to those most in need.

The Pires [27] proposal for increasing the datasets can be used in the future for better training of the network, thus increasing our datasets.

There is a concern with the anonymization of these data, which is a growing concern in the world we live in due to the power of sensitive data on people, a future work is the union of this solution with technologies such as ARM secure zone and others for the protection of this data.

REFERENCES

[1] [n. d.]. Diabetic retinopathy. https://www.nhs.uk/conditions/diabetic-retinopathy/

[2] [n. d.]. Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico,2006. http://bvsms.saude.gov.br/bvs/publicacoes/vigitel_brasil_2006.pdf

[3] [n. d.]. Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico,2017. https://bvsms.saude.gov.br/bvs/publicacoes/vigitel_brasil_2017_vigilancia_fatores_riscos.pdf

[4] 2019. Diabetic Retinopathy. https://afamilyoptician.co.uk/diabetic-retinopathy/

[13]https://www.phelcom.com.br/blog/retinografo-portatil-phelcom-eyer/

[14]https://www.d-eyecare.com/en$_US/shop$

[15]https://www.kaggle.com/reighns/metric-quadratic-weighted-kappa

[5] Mônica de Cássia Alves, José Barreto Carvalheira, Carolina Maria Módulo, and Eduardo Melani Rocha. 2008. Tear film and ocular surface changes in diabetes mellitus. *Arquivos brasileiros de oftalmologia* 71, 6 (2008), 96–103.

[6] Emily Dawn Cole, Eduardo Amorim Novais, Ricardo Noguera Louzada, and Nadia K Waheed. 2016. Contemporary retinal imaging techniques in diabetic retinopathy: a review. *Clinical & Experimental Ophthalmology* 44, 4 (2016), 289–299.

[7] Alcione Brasileiro Oliveira Cunha, Ligia Maria Vieira-da Silva, et al. 2010. Health services accessibility in a city of Northeast Brazil. *Cadernos de Saúde Pública* 26, 4 (2010), 725–737.

[8] John Dilley, Bruce Maggs, Jay Parikh, Harald Prokop, Ramesh Sitaraman, and Bill Weihl. 2002. Globally distributed content delivery. *IEEE Internet Computing* 6, 5 (2002), 50–58.

[9] Mohammad Fadia T., Shaya; Aljawadi. 2007. Diabetic Retinopathy. *Clinical Ophthalmology* 1, 3 (2007), 259–265.

[10] Clément Farabet, Berin Martini, Polina Akselrod, Selçuk Talay, Yann LeCun, and Eugenio Culurciello. 2010. Hardware accelerated convolutional neural networks for synthetic vision systems. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*. IEEE, 257–260.

[11] Donald S Fong, Lloyd Aiello, Thomas W Gardner, George L King, George Blankenship, Jerry D Cavallerano, Fredrick L Ferris, and Ronald Klein. 2004. Retinopathy in diabetes. *Diabetes care* 27, suppl 1 (2004), s84–s87.

[12] Simon S. Haykin. 2009. *Neural networks and learning machines* (third ed.). Pearson Education, Upper Saddle River, NJ.

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. Deep Residual Learning for Image Recognition. *CoRR* abs/1512.03385 (2015). arXiv:1512.03385 http://arxiv.org/abs/1512.03385

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[15] Kaggle and EyePacs. 2015. Kaggle Diabetic Retinopathy Detection. https://www.kaggle.com/c/diabetic-retinopathy-detection/data

[16] Nour Eldeen Khalifa, Mohamed Loey, Mohamed Taha, and Mohamed Taha. 2019. Deep Transfer Learning Models for Medical Diabetic Retinopathy Detection. *Acta Informatica Medica* 27 (12 2019), 327. https://doi.org/10.5455/aim.2019.27.327-332

[17] David B. Kirk and Wen-mei W. Hwu. 2010. *Programming Massively Parallel Processors: A Hands-on Approach* (1st ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[18] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2017. Imagenet classification with deep convolutional neural networks. *Commun. ACM* 60, 6 (2017), 84–90.

[19] David R Lairson, Jacqueline A Pugh, Asha S Kapadia, Ronald J Lorimor, James Jacobson, and Ramon Velez. 1992. Cost-effectiveness of alternative methods for diabetic retinopathy screening. *Diabetes care* 15, 10 (1992), 1369–1377.

[20] GP Leese, S Ahmed, RW Newton, RT Jung, A Ellingford, P Baines, S Roxburgh, and J Coleiro. 1993. Use of mobile screening unit for diabetic retinopathy in rural and urban areas. *British Medical Journal* 306, 6871 (1993), 187–189.

[21] Jose Manuel Santos de Varge Maldonado, Alexandre Barbosa Marques, and Antonio Cruz. 2016. Telemedicine: challenges to dissemination in Brazil. *Cadernos de saude publica* 32 (2016), e00155615.

[22] Diabetes Mellitus. 2005. Diagnosis and classification of diabetes mellitus. *Diabetes care* 28, S37 (2005), S5–S10.

[23] Uwe Meyer-Baese. 2007. *Digital Signal Processing with Field Programmable Gate Arrays* (3rd ed.). Springer Publishing Company, Incorporated.

[24] Varnit Mittal and Bharat Bhushan. 2020. Accelerated Computer Vision Inference with AI on the Edge. In *2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT)*. IEEE, 55–60.

[25] Ana Paula Cavalcante de Oliveira, Mariana Gabriel, Mario Roberto Dal Poz, and Gilles Dussault. 2017. Challenges for ensuring availability and accessibility to health care services under Brazil's Unified Health System (SUS). *Ciência & Saúde Coletiva* 22 (2017), 1165–1180.

[26] Francisco J Pasquel, Andrew M Hendrick, Martha Ryan, Emily Cason, Mohammed K Ali, and KM Venkat Narayan. 2016. Cost-effectiveness of different diabetic retinopathy screening modalities. *Journal of diabetes science and technology* 10, 2 (2016), 301–307.

[27] Ramon Pires, Sandra Avila, Jacques Wainer, Eduardo Valle, Michael D. Abramoff, and Anderson Rocha. 2019. A data-driven approach to referable diabetic retinopathy detection. *Artificial Intelligence in Medicine* 96 (2019), 93 – 106. https://doi.org/10.1016/j.artmed.2019.03.009

[28] Prateek Prasanna, Shubham Jain, Neelakshi Bhagat, and Anant Madabhushi. 2013. Decision support system for detection of diabetic retinopathy using smartphones. In *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*. IEEE, 176–179.

[29] Harry Pratt, Frans Coenen, Deborah M. Broadbent, Simon P. Harding, and Yalin Zheng. 2016. Convolutional Neural Networks for Diabetic Retinopathy. In *20th Conference on Medical Image Understanding and Analysis, MIUA 2016, Loughbrough University, Leicestershire, UK, July 6-8, 2016 (Procedia Computer Science)*, Alastair G. Gale and Yan Chen (Eds.), Vol. 90. Elsevier, 200–205. https:

//doi.org/10.1016/j.procs.2016.07.014

[30] P Vishnu Priya, A Srinivasarao, and JVC Sharma. 2013. Diabetic Retinopathy-Can Lead To Complete Blindness. *Int. J. Sci. Invent. Today* 2, 4 (2013), 254–265.

[31] Dr.Fabio Queiroz. 2018. Campina Grande Dataset. Thisisaprivatedataset

[32] Gwenole Quellec, Loïc Bazin, Guy Cazuguel, Ivan Delafoy, Beatrice Cochener, and Mathieu Lamard. 2016. Suitability of a low-cost, handheld, nonmydriatic retinograph for diabetic retinopathy diagnosis. *Translational vision science & technology* 5, 2 (2016), 16–16.

[33] Ramachandran Rajalakshmi, Radhakrishnan Subashini, Ranjit Mohan Anjana, and Viswanathan Mohan. 2018. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye* 32, 6 (2018), 1138–1144.

[34] Mahadev Satyanarayanan. 2017. The emergence of edge computing. *Computer* 50, 1 (2017), 30–39.

[35] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[36] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2014. Going Deeper with Convolutions. arXiv:cs.CV/1409.4842

[37] Christian Szegedy, Alexander Toshev, and Dumitru Erhan. 2013. Deep Neural Networks for Object Detection. 1–9.

[38] MV Valueva, NN Nagornov, PA Lyakhov, GV Valuev, and NI Chervyakov. 2020. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathematics and Computers in Simulation* (2020).

[39] X. Wang, Y. Lu, Y. Wang, and Wei bang Chen. 2018. Diabetic Retinopathy Stage Classification Using Convolutional Neural Networks. *2018 IEEE International Conference on Information Reuse and Integration (IRI)* (2018), 465–471.

[40] Yingfeng Zheng, Mingguang He, and Nathan Congdon. 2012. The worldwide epidemic of diabetic retinopathy. *Indian journal of ophthalmology* 60, 5 (2012), 428.