

<https://doi.org/10.23913/ride.v11i22.856>

Artículos científicos

Clasificación automática de anastomosis mediante redes neuronales convolucionales en video fetoscópico

An Automatic Classification of Anastomosis by Convolutional Neural Networks in Fetoscopic Video

Classificação automática de anastomoses usando redes neurais convolucionais em vídeo fetoscópico

Ángel Mario Lerma Sánchez

Universidad Da Vinci, México

alerna14@udavinci.edu.mx

<https://orcid.org/0000-0002-5759-1302>

Adriana Mexicano Santoyo

Tecnológico Nacional de México, Instituto Tecnológico de Cd. Victoria, México

mexicanao@gmail.com

<https://orcid.org/0000-0002-8045-7558>

Fabiola Mirosłaba Villalobos Castaldi

Universidad Da Vinci, México

mwillaloboscald@gmail.com

<https://orcid.org/0000-0002-6967-5706>

Pedro Damián Reyes

Universidad de Colima, México

damian@uclm.mx

<https://orcid.org/0000-0003-1114-9612>

Resumen

El síndrome de transfusión gemelo a gemelo (*the twin-twin transfusion syndrome* o *TTTS* por sus siglas en inglés) es el resultado de un flujo sanguíneo desigual a través de anastomosis (conexión de los vasos sanguíneos) vasculares placentarias que enlazan las dos circulaciones fetales. Las anastomosis vasculares en la placenta compartida están presentes en prácticamente todos los embarazos gemelares monocoriónicos (*monochorionic twins* o *MC* por sus siglas en inglés), pero solo en aproximadamente 10 % conducen al síndrome de transfusión gemelo a gemelo. Sin intervención, la condición es a menudo fatal para ambos gemelos. Una alternativa para el tratamiento del *TTTS* es el procedimiento de láser placentario, conocido como *cirugía fetal*, que consiste, de manera general, en dividir la placenta en dos mediante la cauterización láser de los vasos sanguíneos entre los fetos, equilibrando así los flujos sanguíneos.

En la actualidad la cirugía fetoscópica es un procedimiento que se realiza con frecuencia en el país, por lo que la clasificación de anastomosis es vital para esta cirugía, la cual constituye el tratamiento más recomendado. Por su grado de complejidad, esta intervención quirúrgica presenta múltiples dificultades: los fetos se mueven durante el procedimiento, la orientación del video empleado no es adecuada para un análisis más preciso y el campo de visión generado por el fetoscópio es muy pequeño; por lo tanto, es necesario contar con una herramienta que ayude al médico novel a poder diferenciar y clasificar las anastomosis de una manera más adecuada. El objetivo de este trabajo, por tanto, es presentar el desarrollo de una herramienta computacional que contribuya a la clasificación automática de anastomosis dentro de un video fetoscópico mediante redes neuronales convolucionales de manera que sirva de soporte para el médico novel en su etapa de entrenamiento.

Para ello, se construyó un DataSet (conjunto de imágenes) a partir de videos fetoscópicos sin clasificar que luego se catalogaron en tres categorías, seleccionadas conjuntamente con los expertos, mediante una herramienta computacional *ex profeso* para ello (VideoLabel). La técnica de aumento de datos sirvió para construir imágenes artificiales a partir de las reales ya clasificadas, dado que el número de imágenes etiquetadas no eran suficientes; en ese contexto, se seleccionó la arquitectura AlexNet para realizar una transferencia de aprendizaje y para entrenarse, lo que arrojó resultados con una efectividad superior a 90 %. Estos datos permiten concluir que de no contar con una herramienta que permitiera entrenar a los médicos noveles en la identificación de anastomosis, su capacitación

tomaría más tiempo en llevarse a cabo, ya que se realizaría de manera *in situ* en cada cirugía fetal. Sin embargo, en esta investigación se generó el diseño de un *software* con el que resulta factible la clasificación automática de anastomosis a partir de un video fetoscópico mediante una red neuronal convolucional, con resultados prometedores como herramienta de apoyo para los médicos en su etapa de entrenamiento.

Palabras clave: aprendizaje de máquina, aprendizaje supervisado, inteligencia artificial, síndrome de transfusión gemelo a gemelo.

Abstract

Twin-twin transfusion syndrome (*TTTS*) is the result of uneven blood flow through placental vascular anastomosis (blood vessel connection) that link the two fetal circulations. Vascular anastomosis in the shared placenta are present in virtually all monochorionic twin pregnancies (MCs), but in only about 10% lead to twin-twin transfusion syndrome. Without intervention, the condition is often fatal for both twins. An alternative to *TTTS* treatment is the placental laser procedure known as fetal surgery, which consists, in a very general way, of splitting the placenta in two, by laser cauterization of blood vessels between fetuses, thus balancing blood flows.

Currently fetoscopic surgery is a procedure that is performed frequently in Mexico and the appropriate classification of anastomosis is vital for this surgery, since it represents the most recommended treatment. However, due to its degree of complexity, this surgical intervention presents multiple difficulties, such as fetuses moving during the procedure, the orientation of the video used is not suitable for a more accurate analysis, the field of view generated by the fetus is very small. Therefore, it is necessary to have a tool that helps the doctor to be able to differentiate and classify anastomosis in a more appropriate way. The objective of this work is to present the development of a computational tool that contributes to the automatic classification of anastomosis within a fetoscopic video through convolutional neural networks in a way that serves as a support for the unexperienced doctor in his training stage.

A DataSet (image set) was built from fetoscopic videos first unclassified, then cataloged into three categories, selected in conjunction with experts, using a computational tool specifically created for this purpose (VideoLabel). The data augmentation technique served to build artificial images from the actual ones already classified, since the number of

tagged images were not sufficient; in the same context, the AlexNet architecture was selected to perform a learning transfer and be trained to obtain results above 90% effectiveness in the classifications made by the computational tool created. This data allows us to conclude that if we did not have a tool that would allow first-time physicians to train in the identification of anastomosis as a practice prior to fetal surgery, their training will take longer since it is done in situ in each fetal surgery. As a result of this research, the design of a software was generated with which it is feasible to automatically classify anastomosis from a fetoscopic video through a convolutional neural network with promising results as a tool to support doctors in their training stage, allowing them to carry out their training in a more agile and less timely way.

Keywords: Machine Learning, Supervised Learning, Artificial Intelligence, Twin to Twin Transfusion Syndrome.

Resumo

A síndrome da transfusão de gêmeos (TTTS) é o resultado do fluxo sanguíneo desigual através das anastomoses vasculares da placenta (conexão dos vasos sanguíneos) ligando as duas circulações fetais. Anastomoses vasculares na placenta compartilhada estão presentes em praticamente todos os gêmeos monócóricos (MCs), mas apenas em cerca de 10% levam à síndrome de transfusão de gêmeos. Sem intervenção, a condição costuma ser fatal para ambos os gêmeos. Uma alternativa para o tratamento do TTTS é o procedimento a laser placentário, conhecido como cirurgia fetal, que geralmente consiste em dividir a placenta em duas por meio da cauterização a laser dos vasos sanguíneos entre os fetos, equilibrando assim os fluxos sanguíneos.

Atualmente, a cirurgia fetoscópica é um procedimento muito realizado no país, portanto a classificação da anastomose é fundamental para esta cirurgia, que é o tratamento mais recomendado. Devido ao seu grau de complexidade, esta intervenção cirúrgica apresenta múltiplas dificuldades: os fetos movimentam-se durante o procedimento, a orientação do vídeo utilizado não é adequada para uma análise mais precisa e o campo de visão gerado pelo fetoscópio é muito pequeno; portanto, é necessário um instrumento que ajude o médico novato a diferenciar e classificar as anastomoses de forma mais adequada. O objetivo deste trabalho, portanto, é apresentar o desenvolvimento de uma ferramenta computacional que contribua para a classificação automática de anastomose em um vídeo fetoscópico por meio

de redes neurais convolucionais de forma a servir de suporte para o médico iniciante em seu estágio de treinamento.

Para isso, um DataSet (conjunto de imagens) foi construído a partir de vídeos fetoscópicos não classificados que foram então catalogados em três categorias, selecionadas em conjunto com os especialistas, usando uma ferramenta computacional especialmente projetada (VideoLabel). A técnica de aumento de dados serviu para construir imagens artificiais a partir das reais já classificadas, uma vez que o número de imagens rotuladas não era suficiente; Neste contexto, a arquitetura AlexNet foi selecionada para realizar uma transferência de aprendizagem e para treinamento, que rendeu resultados com uma eficácia superior a 90%. Esses dados permitem concluir que, na ausência de uma ferramenta que possibilitasse o treinamento de novos médicos na identificação de anastomoses, o seu treinamento demoraria mais, pois seria realizado in loco em cada cirurgia fetal. No entanto, esta pesquisa gerou o desenho de um software com o qual é viável a classificação automática de anastomoses de um vídeo fetoscópico utilizando uma rede neural convolucional, com resultados promissores como ferramenta de apoio a médicos em estágio de formação.

Palavras-chave: aprendizado de máquina, aprendizado supervisionado, inteligência artificial, síndrome de transfusão de gêmeos.

Fecha Recepción: Julio 2020

Fecha Aceptación: Febrero 2021

Introduction

Twin-twin transfusion syndrome (TTTS) is the result of uneven blood flow through placental vascular anastomoses (which is the connection of blood vessels) that link the two fetal circulations. These shared placental vascular anastomoses are present in virtually all monozygotic twins (MZs), but only in about 10% lead to twin-to-twin transfusion syndrome.

Net blood transfusion occurs at the expense of the so-called donor twin, which becomes hypovolemic and anemic, while the recipient twin becomes hypervolemic and polycythemic. This leads to the rapid development of a significant mismatch in amniotic fluid volume between twins that without proper intervention can be fatal for both twins (Spruijt, Lopriore, Steggerda, Slaghekke, & van Klink, 2019).

Twin-to-twin transfusion syndrome is one of the deadliest circumstances in fetal medicine today, and it remains a great challenge for obstetricians and neonatologists around

the world. In this sense, the best treatment is fetoscopic laser surgery. The technique, currently in force, uses such a laser to interrupt blood flow through vascular communications in order to coagulate only the anastomotic vessels in the vascular equator that is at a certain distance from the interlaced membrane due to the discordant volumes of amniotic fluid between twins (Glennon, Shemer, Palma-Dias y Umstad, 2016).

Laser surgery for twin-to-twin transfusion syndrome can be divided into two fundamental steps: 1) endoscopic identification of the placental vascular anastomoses, and 2) laser ablation of the anastomoses. Therefore, the goal of laser surgery is to correctly identify and remove all vascular anastomoses. This raises two questions: 1) can all placental vascular anastomoses be identified? And 2) can all placental vascular anastomoses be removed? (Quintero, Kontopoulos and Chmait, 2016).

Now, focusing on endoscopic detection of placental vascular anastomoses, the following question arises: how can the anastomoses be identified? Some authors suggested that anastomoses could be distinguished according to their appearance, that is, according to specific patterns, angles, with drawings intended to help in this process, that is, identification based on pattern recognition. Unfortunately, this was not possible for clinicians due to the large number of patterns that anastomoses can have (De Lia, Kuhlmann, Cruikshank & O'Bee, 1993).

From the above, the following research questions arise: can a computational tool help in the identification of anastomoses within a fetoscopic video, and how accurate can this identification be? The present work hypothesizes that a computational technique can identify anastomoses with a certain percentage of accuracy.

Distinguishing the anastomoses to perform an adequate intervention in laser surgery for twin-twin transfusion syndrome constitutes a problem for the less experienced physician, therefore, in order to meet this need, the following objectives are set: to demonstrate that a tool Computational analysis can help in the identification of anastomosis in a fetoscopic video with a certain percentage of accuracy, and create a computational tool that helps in the automatic classification of anastomoses, in such a way that it serves the novice doctor as a support tool in their training and habilitation stage.

The science of solving clinical problems, through the analysis of images generated in clinical practice, is known as medical image analysis, and its objective is to extract information efficiently to improve clinical diagnosis. This article highlights this advance,

which consists of the application of machine learning techniques for the analysis of medical images. It presents, in the first section, a description of the development of the research work carried out with several sections that give clarity about the process followed in order to achieve the planned objectives, highlighting the way in which deep learning is used successfully as a tool to machine learning, where a neural network is able to learn characteristics automatically. Of the deep learning techniques, deep convolutional networks are actively used for medical image analysis. This includes application areas such as segmentation, abnormality detection, disease classification, computer-aided diagnosis, and recovery (Anwar et al., 2018). But what do you mean when you talk about convolutional neural networks?

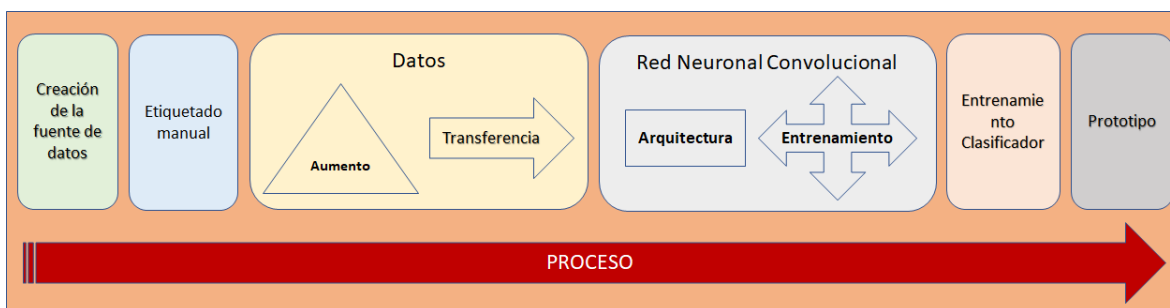
A convolutional neural network is a type of deep learning model for processing data that has a grid model - such as images - inspired by the organization of the animal visual cortex and designed to automatically and adaptively learn spatial hierarchies of features, from low to high standard levels. A convolutional neural is a mathematical construct that is generally made up of three types of layers (or building blocks): convolution, clustering, and fully connected layers. The first two perform feature extraction, while the third — a fully connected layer — maps the extracted features to the final output, such as classification.

A convolution layer plays a key role in the convolutional neural, which is made up of a stack of mathematical operations, such as convolution, a specialized type of linear operation. In digital images, pixel values are stored in a two-dimensional (two-dimensional) grid, that is, an array of numbers, and a small parameter grid called the kernel, a tunable feature extractor, is applied at each image position. , which makes convolutional neurals highly efficient for image processing, as a feature can appear anywhere in the image. As one layer feeds its output to the next layer, the extracted features can become hierarchical and progressively more complex. The process of optimizing parameters - such as cores - is called training, which is performed to minimize the difference between the outputs and the terrain truth labels through an optimization algorithm called backward propagation and gradient descent, among others. (Yamashita, Nishio, Do y Togashi, 2018).

Method

This section explains the methodology that was followed to propose a way to solve the problem presented, from the collection of the input data, its treatment and transformation to the proposal of a tool that helps to solve it; Figure 2 shows the methodology followed in graphic form.

Figura 1. Metodología seguida



Fuente: Elaboración propia

Create the data source

In the context of this work, the general process of the data origin provided were videos in MPG format, recorded during the surgical intervention to real patients when fetal surgery was applied to them; The videos were extracted directly from the fetoscope by the doctor in charge of the intervention and include the following data: date, treatment, surgeon, patient ID, date of birth and gender, in an XML or HTML file. They are generally N videos with a structure named Video_001.mpg,... Video_00N.mpg for each patient; For this research, N videos of 12 patients were accessed.

Subsequently, a program was created to join all the separate videos of each patient (12 complete videos were obtained from this process), in this way it was possible to have them in MP4 format for later use of anastomosis classification with the Videolabel tool by the doctor expert.

In addition to this, from the 12 complete videos the frames of each one were extracted and renamed with a suitable nomenclature to have an unclassified image dataset. The nomenclature was as follows:

$F = \text{Frame.}$

$P1, P2, P3 \dots PN = \text{Patient.}$

0000000001 = Consecutive number.

Example:

$F-P1-0000000001.png$

To conclude, in this phase a program was developed in MatLab, where the entire directory containing images was reviewed and they were automatically cropped in order to leave only the image itself and eliminate the excess. This activity was carried out because it works with thousands of images, which will have to be processed by the convolutional neural network. This segment removal task saves time and results in a DataSet of unclassified images.

Manual labeling

For the labeling of the image DataSet, software was developed that helps to classify the images (called VideoLabel). This classifier software was used by medical experts with knowledge of the subject; its operation, in general, consists of reading an already processed video. The software allows an expert doctor -in case of identifying an anastomosis- to pause the video, mark the item and save the type of anastomosis identified while internally the program saves said classified image. The classes of anastomosis that the VideoLabel program already has loaded are the following: normal, arterio-venous, veno-arterial, arterio-arterial, veno-venous, placental border, interfetal membrane, insertion of the donor twin cord, insertion of the cord of the receptor twin, capillary.

The output of the VideoLabel program is the storage of the classified images in a folder that make up the DataSet of images labeled by the expert. In this research work we worked with the arterio-venous, capillary and veno-arterial categories; Figure 2 shows the three types of anastomoses with which we worked:

Figura 2. Categorías de imágenes del trabajo de investigación



Fuente: Elaboración propia

It should be noted that not all types of anastomoses are present in the interventions; likewise, it is difficult to assemble a large image DataSet, since the time spent on manual labeling by the expert physician is limited. The next section specifies the quantities of images collected and tagged by the expert.

Data augmentation

In general, convolutional neural networks are too deep to train from scratch with small data sets (you need at least a thousand images classified by category) for training from scratch. For that, learning transfer techniques and data augmentation are recommended; In this sense, the data augmentation technique resides in artificially increasing the group of training images using various readjustments to the original images, such as modifying the brightness, scaling, zooming in, rotating, reflecting vertically and horizontally, etc. The images obtained must be as acceptably real as possible so that an external viewer cannot distinguish between an image created by increased data and an original one (Gómez-Ros et al., 2019).

Although there are various strategies for data augmentation, the previous ones have been explored only with natural images, and not with medical images. Among the strategies for increasing data in medical images, the following are recommended, since they were the ones that gave the best results, since they increased the percentage of effectiveness in the training of convolutional neural networks: flips (turn around), rotate (rotation), gaussian blur (gaussian blur) (Hussain, Gimenez, Yi y Rubin, 2018).

In this investigation, 2320 artificial images were created to increase the number of total images per category; Table 1 shows the set of real and artificial images that make up the DataSet.

Tabla 1. Resumen de imágenes clasificadas y generadas

CLASIFICACIÓN	REALES	FLIP	ROT	BLUR	TOTAL
Arterio-venosa	72	216	792	144	1224
Capilar	29	87	319	58	493
Veno-arterial	44	132	484	88	748
Total	145	435	1595	290	2465

Fuente: Elaboración propia

Learning transfer

Learning transfer is a common and effective strategy for training a network with a small data set, where there is a pre-trained network on an extremely large data set, such as AlexNet, which contains 1.2 million images with 1000 classes, which are then reuse and apply to the given task of interest. The assumption underlying the transfer of learning is that generic characteristics (that is, what is learned in a large enough data set) can be shared between seemingly disparate data sets. This portability of learned generic features is a unique advantage of deep learning that is useful in various domain tasks with small data sets (Yamashita et al., 2018). Similarly, learning transfer from the pre-trained AlexNet network was performed, where the last layers were adjusted (this is explained in the following sections of this article).

Convolutional neural network architecture

For some years now, various models of deep convolutional neural networks have been exposed, such as AlexNet, ZFNet, VGG, GoogleNet, Residual Net, DenseNet, FCN and U-Net. Such models have not only provided state-of-the-art performance for various tasks such as image classification, segmentation, object detection, and tracking tasks, but have also offered a new point of view for image fusion (Zhou, Ruan and Canu, 2019).

The AlexNet pre-trained network has been used for this research work —as suggested in Tajbakhsh et al. (2016) -, since it is considered adequate to carry out learning transfer in medical images. Also, as the data sets are still too small, instead of adjusting all the layer weights, two fully connected layers have been added to the end of the previous layers, for which a layer was added to the network architecture convolutional neuronal. This layer is called fully connected layer (23 in the list of layers for AlexNet), which is in charge of classifying the data in several classes where the layers before it (from 1 to 22) perform the task of extracting the characteristics of the images.

Regarding the classification output layer (25 in the AlexNet layer list), they are in charge of calculating the loss of cross entropy for classification problems of multiple classes with mutually exclusive classes. The layer infers the number of classes from the output size of the previous layer.

It should be noted that the transfer of learning helps to train the two layers that have been modified according to the needs, and not the entire network; Figure 3 shows the final layer architecture of the convolutional neural network.

Figura 3. Capas de la red neuronal convolucional

1	data 227x227x3 images with 'zerocenter' normalization	Image Input	227x227x3	-	0
2	conv1 96 11x11x3 convolutions with stride [4 4] and padding [0 0 0 0]	Convolution	55x55x96	Weights 11x11x3x96 Bias 1x1x96	34944
3	relu1 ReLU	ReLU	55x55x96	-	0
4	norm1 cross channel normalization with 5 channels per element	Cross Channel Nor...	55x55x96	-	0
5	pool1 3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	27x27x96	-	0
6	conv2 256 5x5x48 convolutions with stride [1 1] and padding [2 2 2 2]	Convolution	27x27x256	Weights 5x5x48x256 Bias 1x1x256	307456
7	relu2 ReLU	ReLU	27x27x256	-	0
8	norm2 cross channel normalization with 5 channels per element	Cross Channel Nor...	27x27x256	-	0
9	pool2 3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	13x13x256	-	0
10	conv3 384 3x3x256 convolutions with stride [1 1] and padding [1 1 1 1]	Convolution	13x13x384	Weigh... 3x3x256x3... Bias 1x1x384	885120
11	relu3 ReLU	ReLU	13x13x384	-	0
12	conv4 384 3x3x192 convolutions with stride [1 1] and padding [1 1 1 1]	Convolution	13x13x384	Weigh... 3x3x192x3... Bias 1x1x384	663936
13	relu4 ReLU	ReLU	13x13x384	-	0
14	conv5 256 3x3x192 convolutions with stride [1 1] and padding [1 1 1 1]	Convolution	13x13x256	Weigh... 3x3x192x2... Bias 1x1x256	442624
15	relu5 ReLU	ReLU	13x13x256	-	0
16	pool5 3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	6x6x256	-	0
17	fc6 4096 fully connected layer	Fully Connected	1x1x4096	Weights 4096x9216 Bias 4096x1	37752832
18	relu6 ReLU	ReLU	1x1x4096	-	0
19	drop6 50% dropout	Dropout	1x1x4096	-	0
20	fc7 4096 fully connected layer	Fully Connected	1x1x4096	Weights 4096x4096 Bias 4096x1	16781312
21	relu7 ReLU	ReLU	1x1x4096	-	0
22	drop7 50% dropout	Dropout	1x1x4096	-	0
23	fc 3 fully connected layer	Fully Connected	1x1x3	Weights 3x4096 Bias 3x1	12291
24	prob softmax	Softmax	1x1x3	-	0
25	classoutput crossentropyx with 'Arterio_venosa' and 2 other classes	Classification Output	-	-	0

Fuente: Elaboración propia

Convolutional Neural Network Training

For the realization of both the convolutional neural network training and the tests in general, the technical characteristics of the computer equipment used are presented: Intel Core i5 7200U processor at 2.5Ghz, 8 Gb RAM, Windows 10 operating system of 64 bits, and Nvidia GeForce 940mx 2 Gb RAM graphics card.

The total number of images (between real and artificial) to train the convolutional neural network was 2465. In this regard, it is important to note that to test the network once trained, 23 unpublished images were used (that is, the network did not know them) .

The activities followed to proceed with the training of a convolutional neural network are listed below:

1. Initialize all the parameters or weights with random values of the first layers of the network.
2. Take a set of training images and use them on the model.
3. Compute the total error of the probabilities as a consequence of the model.
4. Extend back to compute the gradient error of all the weights in the network and use the descending gradient to modify these values and decrease the output error (Quintero, Merchán, Cornejo y Sánchez Galán, 2018).

An important parameter refers to the number of epochs that the convolutional neural network will have; an epoch refers to a single pass of the complete training set (Schilling, 2016). Knowing the number of epochs to train the convolutional neural network in medical images is an important factor to have a decision parameter; some authors recommend 50 epochs (Kumar, Kim, Lyndon, Fulham and Feng, 2017; Petscharnig and Schöffmann, 2018).

Figure 4 shows the convolutional neural network performing the training for a maximum of 50 epochs. As a result, an accuracy percentage of 100% was obtained and the time used by the program was approximately 2 hours and 21 minutes.

Figura 4. Red neuronal convolucional entrenándose

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:06	12.50%	2.6806	0.0010
1	50	00:01:24	29.69%	1.8023	0.0010
1	100	00:02:45	29.69%	1.7946	0.0010
2	150	00:04:10	51.56%	1.5030	0.0010
2	200	00:05:29	48.44%	1.2927	0.0010
3	250	00:06:50	59.38%	1.1698	0.0010
3	300	00:08:06	64.06%	1.0036	0.0010
4	350	00:09:23	67.19%	0.8340	0.0010
4	400	00:10:46	85.94%	0.5112	0.0010
5	450	00:12:04	87.50%	0.4826	0.0010
5	500	00:13:27	73.44%	0.6337	0.0010
•					
•					
•					
46	4550	02:12:12	100.00%	0.0045	0.0010
46	4600	02:13:16	98.44%	0.0130	0.0010
47	4650	02:14:19	100.00%	0.0042	0.0010
47	4700	02:15:22	100.00%	0.0023	0.0010
48	4750	02:16:26	100.00%	0.0018	0.0010
48	4800	02:17:29	100.00%	0.0137	0.0010
49	4850	02:18:32	100.00%	0.0123	0.0010
49	4900	02:19:35	100.00%	0.0022	0.0010
50	4950	02:20:40	100.00%	0.0195	0.0010
50	5000	02:21:42	100.00%	0.0036	0.0010

Fuente: Elaboración propia

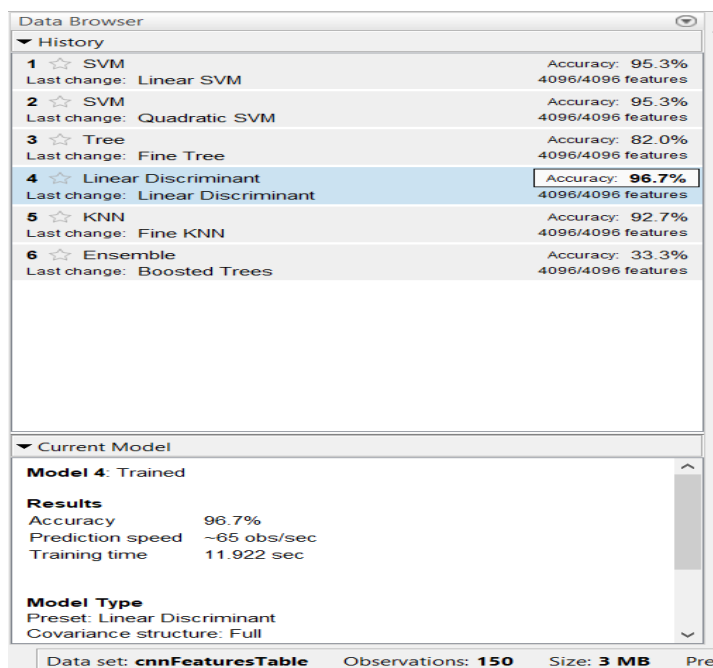
Training a classifier

Cataloging an object consists of naming it to one of the available classes. Objects can be determined by a list of particulars, such as the texture, size or color of their pixels; to be able to catalog objects it is necessary to specify the limits between the different classes. Typically these limits are computed through a training treatment using the particularities of a series of example models of the classes. Limits are mentioned for clarity; in general, a classifier deduces decision patterns during training.

From this it follows that cataloging an unknown object consists of designating it to the class in which the particularities used during the training are more equivalent to the particularities of the object (Garrido Satué, 2013).

Once the training of the convolutional neural network has been completed, the training of a classifier is applied to the result of said training, since the data can be explored, selected particularities, specify cross-validation schemes, train models and evaluate results. . Automatic training is likely to be performed to find out the preferable type of classifier model (such as decision trees, discriminant analysis, support vector machines, logistic regression, close neighbors, assembled classification). Figure 5 shows the selected classifier with data from various trained classifiers and which one obtains the best classifier model for the data presented.

Figura 5. Clasificador seleccionado



Data Browser		
▼ History		
1 ☆ SVM	Accuracy: 95.3%	4096/4096 features
Last change: Linear SVM		
2 ☆ SVM	Accuracy: 95.3%	4096/4096 features
Last change: Quadratic SVM		
3 ☆ Tree	Accuracy: 82.0%	4096/4096 features
Last change: Fine Tree		
4 ☆ Linear Discriminant	Accuracy: 96.7%	4096/4096 features
Last change: Linear Discriminant		
5 ☆ KNN	Accuracy: 92.7%	4096/4096 features
Last change: Fine KNN		
6 ☆ Ensemble	Accuracy: 33.3%	4096/4096 features
Last change: Boosted Trees		
▼ Current Model		
Model 4: Trained		
Results		
Accuracy	96.7%	
Prediction speed	~65 obs/sec	
Training time	11.922 sec	
Model Type		
Preset: Linear Discriminant		
Covariance structure: Full		
Data set: cnnFeaturesTable Observations: 150 Size: 3 MB Pre		

Fuente: Elaboración propia

Prototype operation

A prototype was developed to test the operation of the convolutional neural network, which was done with an end user who did not have much computer knowledge. The screen has three action buttons: browse, pause and exit. In Figure 6 you can see a screen of the prototype working in the anastomosis classification.

The data source must be a folder that contains the video of some surgical intervention in MP4 format; Subsequently, the prototype will reproduce the video automatically while the

convolutional neural network selects each frame or frame, analyzes it, and when finding a match with the trained categories, it will mark them with a color (green for arterio-venous, yellow for capillary and blue for veno-arterial); at the end it will save said frame in the device. The nomenclature with which the image is saved is as follows, according to the selected categories:

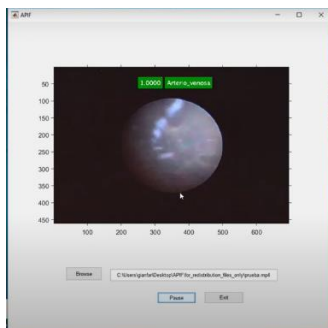
AV-fecha-hora.png for category *arterio-venosa*.

C-fecha-hora.png for category *capilar*.

VA-fecha-hora.png for category *veno-arterial*.

The prototype output is a set of images classified and labeled by the convolutional neural network, according to the three previously mentioned categories. With the use of the prototype, the specialist doctor can have results to validate them.

Figura 6. Muestra de funcionamiento del prototipo de *software*



Fuente: Elaboración propia

Results

The main scope of this research work was to determine if a computational tool could help identify anastomoses within a fetoscopic video and how accurate that identification would be. In this sense, during the development some results of the work carried out were obtained. For example:

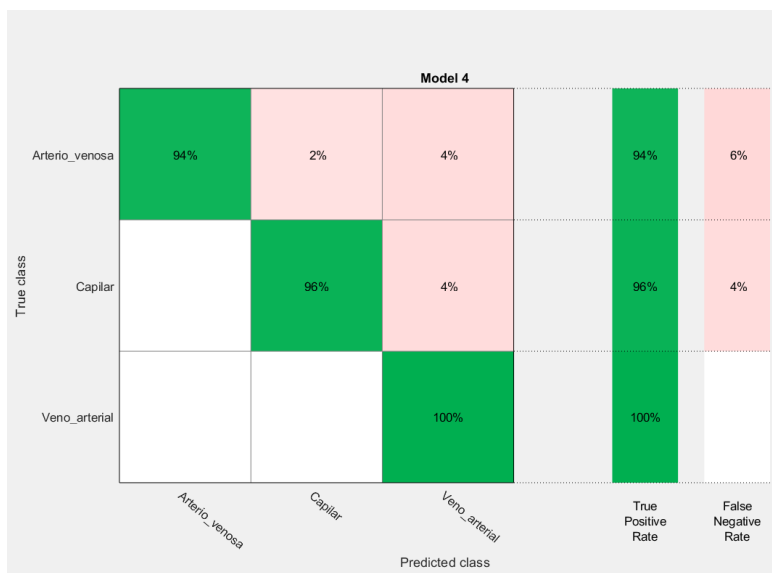
- 12 complete stitched videos of the surgeries were obtained with more than eight hours of fetoscopic video.
- From the 12 videos, a DataSet of 111 273 unclassified images was obtained.
- Product of manual labeling by the medical expert, the images classified by categories shown in table 1 were obtained.

An important part of the results allows us to maintain that the proposed solution works adequately for what it was designed, for which it was considered essential to apply techniques such as the confusion matrix, the ROC curve and the AUC to indicate whether said methodology is acceptable.

In this sense, and in a general way, it can be mentioned that a confusion matrix is a tool that is commonly used to delineate the performance of a classification system in a cluster of data in which they are known as true values. The matrix includes information from the actual and predicted classifications made by the system, and is used to compute the performance of the algorithm. Another way to examine the performance of the classifiers is the ROC (receiver operating characteristics) graphs, a method to visualize, organize and select the classifiers based on their performance, where the ROC curve contains all the information from the confusion matrix (Fos Guarinos, 2016). Likewise, the area under the AUC curve (area under curve) is the statistic par excellence to measure the discriminant capacity of the test and to compare tests with each other and determine which is the most efficient (Del Valle Benavides, 2017).

In figure 7 you can see the confusion matrix resulting from the experimentation carried out with three different classifications in the convolutional neural network after training the network. Note that for the arterio-venous classification there are 94% true positives and 6% false negatives in the set of test values provided to the convolutional neural network. For the capillary classification, there is 96% true positive and 4% false negative classification in the set of test values supplied to the convolutional neural network. Finally, for the veno-arterial classification there are 100% true positives and 0% false negatives in the set of test values provided to the convolutional neural network.

Figura 7. Matriz de confusión resultante



Fuente: Elaboración propia

Regarding the ROC curve, the following results were achieved for each of the three classifications (Table 2 shows the AUC obtained for each category).

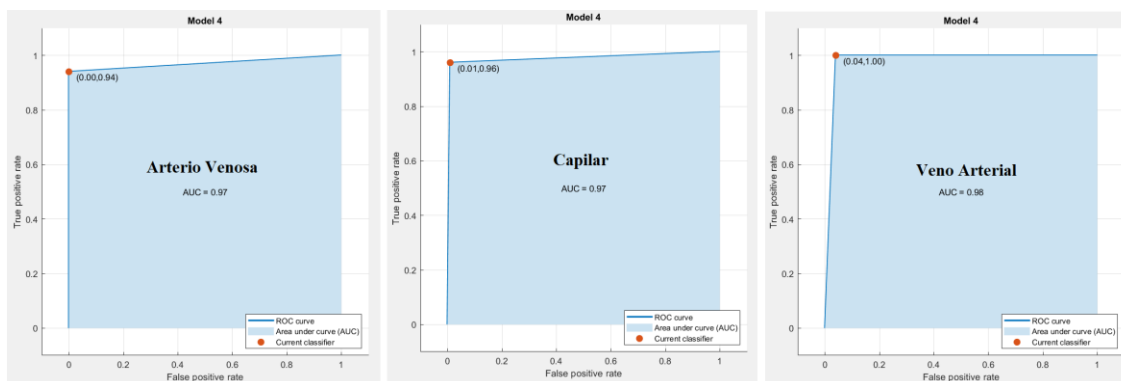
Tabla 2. Resultados de AUC por clasificación

CLASIFICACION	AUC
Arterio-venosa	0.97
Capilar	0.97
Veno-arterial	0.98

Fuente: Elaboración propia

A result of the ROC and its AUC whose value is between 0.5 and 1 indicates that the prediction of the classifier is correct (if it tends to 1 it is almost perfect); on the contrary, a result below 0.5 indicates that there are inconsistencies in the classification of the data. Figure 8 shows each of the categories with its ROC curve and AUC.

Figura 8. Categorías con su curva ROC y AUC



Fuente: Elaboración propia

Discussion

As mentioned in the results, effectiveness percentages greater than 90% were achieved overall for the three categories, with an actual image rate of about 150. A limitation and area of weakness of the research was the relatively low number of images real that was solved with the data augmentation technique; but what if the number of real images increased? Venturing a hypothesis, an increase in the rate of effectiveness can be assumed.

Notwithstanding the above, the great strength of the research was the effectiveness percentages theoretically obtained by the aforementioned techniques, which could be demonstrated with the software prototype created. Consequently, it can be affirmed that the proposed objective was achieved, that is, to develop a computational tool to contribute to the automatic classification of anastomosis within a fetoscopic video using convolutional neural networks, which serve as support in the training stage of the patient. new doctor.

Conclusions

In the present work, the creation of a computational tool to help the novice doctor in the identification of anastomosis within a fetoscopic video was explained, which is part of his preparation to perform an adequate intervention with laser surgery for twin transfusion syndrome to twin.

In this sense, it should be noted that the creation of a prototype that reads and reproduces fetoscopic videos contains within itself a convolutional neural network duly trained with three different categories of anastomosis images, which allows them to be

classified with an effectiveness rate greater than 90%. while the video is playing. However, it is worth noting that the fluidity capacity of the prototype is directly linked to the computer equipment used, so one with high performance is suggested (for example, with a solid state hard disk of at least 500 Gb, at least 8 GB of RAM and a 64-bit operating system).

As an additional product, two image DataSets were obtained: one unclassified and the other duly classified by the doctor or expert, which was the basis with which the convolutional neural network used in the generated tool was trained. Therefore, it can be assured that the prototype created will be of great help for new medical personnel in training to perform fetoscopic surgeries, since it will allow them to have a support tool.

Future lines of research

Finally, as future lines of research, a comparative study and tests with other convolutional neural network architectures (apart from AlexNet), such as VGGNet, GoogleNet, ResNet and ZFNet, can be carried out in order to compare their results, increase the number of categories for make the prototype more complete, increase the number of images in the categories to optimize the accuracy of the results, and apply unsupervised learning to classify videos in order to make adjustments to the network and strengthen it.

Acknowledgment

This work is part of Ángel Mario Lerma Sánchez's doctoral thesis for his doctorate in Computer Systems. It was directed by Dr. Adriana Mexicano Santoyo (advisor); Dra. Fabiola Miroslaba Villalobos Castaldi, co-director; and Dr. Pedro Damián Reyes, co-director. Likewise, we thank Dr. Rogelio Cruz-Martínez for all the facilities for the project and Dr. María Belén Nieto Castro for her hard work in the classification of images and feedback. To Dr. José Ruiz Pinales for his valuable support with the VideoLabel tool. Likewise, thanks are expressed to Da Vinci University for all the support provided.

References

- Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M. and Khan, M. K. (2018). Medical Image Analysis using Convolutional Neural Networks: A Review. *J Med Syst*, 42(11). Doi: 10.1007/s10916-018-1088-1
- De Lia, J. E., Kuhlmann, R. S., Cruikshank, D. P. and O'Bee, L. R. (1993). Current topic: Placental surgery: a new frontier. *Placenta*, 14(5), 477-485. Doi: 10.1016/s0143-4004(05)80201-2
- Del Valle Benavides, A. (2017). *Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones* (tesis de pregrado). Sevilla, España: Universidad de Sevilla. Recuperado de <https://idus.us.es/bitstream/handle/11441/63201/Valle%20Benavides%20Ana%20Roc%20C3%ADo%20del%20TFG.pdf?sequence=1>
- Fos Guarinos, B. (2016). *Diseño de técnicas de inteligencia artificial aplicadas a imágenes médicas de rayos X para la detección de estructuras anatómicas de los pulmones y sus alteraciones* (tesis de pregrado). Valencia, España: Universidad Politécnica de Valencia.
- Garrido Satué, M. (2013). *Reconocimiento de señales de tráfico para un sistema de ayuda a la conducción* (tesis de maestría). Sevilla, España: Universidad de Sevilla, Escuela Técnica Superior de Ingeniería. Recuperado de https://riunet.upv.es/bitstream/handle/10251/70103/73657675Z_TFG_14733814699245714722695301721227.pdf?sequence=2&isAllowed=y
- Glennon, C. L., Shemer, S. A., Palma-Dias, R. and Umstad, M. P. (2016). The History of Treatment of Twin-to-Twin Transfusion Syndrome. *Twin Research and Human Genetics*, 19(3), 168-174. Doi: 10.1017/thg.2016.27
- Gómez-Ros, A., Tabik, S., Luengo, J., Herrera, F., Shihavuddin, A. and Krawczyk, B. (2019). Redes neuronales convolucionales para una clasificación precisa de imágenes de corales. *XVIII Conferencia de la Asociación Española para la Inteligencia Artificial*, 1171-1176. Recuperado de https://backend.orbit.dtu.dk/ws/files/166595939/CAEPIA2018_paper_212.pdf
- Hussain, Z., Gimenez, F., Yi, D. and Rubin, D. (2018). Differential Data Augmentation Techniques for Medical Imaging Classification Tasks. *AMIA Annu Symp Proc.*, 2017,

- 979–984. Retrieved from
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5977656/pdf/2730723.pdf>
- Kumar, A., Kim, J., Lyndon, D., Fulham, M. and Feng, D. (2017). An Ensemble of Fine-Tuned Convolutional Neural Networks for Medical Image Classification. *IEEE Journal of Biomedical and Health Informatics*, 21(1), 31-40. Doi: 10.1109/JBHI.2016.2635663
- Petscharnig, S. and Schöffmann, K. (2018). Learning laparoscopic video shot classification for gynecological surgery. *Multimedia Tools and Applications*, 77, 8061–8079. Doi: <https://doi.org/10.1007/s11042-017-4699-5>
- Quintero, C., Merchán, F., Cornejo, A. y Sánchez Galán, J. (2018). Uso de redes neuronales convolucionales para el reconocimiento automático de imágenes de macroinvertebrados para el biomonitorio participativo. *KnE Engineering*, 3, 585-596. Recuperado de <https://knepublishing.com/index.php/KnE-Engineering/article/view/1462/3527>
- Quintero, R. A., Kontopoulos, E. and Chmait, R. H. (2016). Laser Treatment of Twin-to-Twin Transfusion Syndrome. *Twin Research and Human Genetics*, 19(3), 197-206. Doi: <https://doi.org/10.1017/thg.2016.28>
- Schilling, F. (2016). *The Effect of Batch Normalization on Deep Convolutional Neural Networks* (master thesis). Stockholm, Sweden: KTH Royal Institute of Technology School of Computer Science and Communication. Retrieved from <https://kth.diva-portal.org/smash/get/diva2:955562/FULLTEXT01.pdf>
- Spruijt, M. S., Lopriore, E., Steggerda, S. J., Slaghekke, F. and van Klink, J. M. (2019). Twin-twin transfusion syndrome in the era of fetoscopic laser surgery: antenatal management, neonatal outcome and beyond. *Expert Review of Hematology*, 13(3), 259-267. Doi: <https://doi.org/10.1080/17474086.2020.1720643>
- Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B. and Liang, J. (2016). Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Transactions on Medical Imaging*, 35(5), 1299-1312. Doi: <https://doi.org/10.1109/tmi.2016.2535302>
- Yamashita, R., Nishio, M., Do, R. and Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights Imaging*, 9, 611–629. Doi: <https://doi.org/10.1007/s13244-018-0639-9>

Zhou, T., Ruan, S. and Canu, S. (2019). A review: Deep learning for medical image segmentation using multi-modality fusion. *Array*, 3-4, 100004. Doi: <https://doi.org/10.1016/j.array.2019.100004>

Rol de Contribución	Autor (es)
Conceptualización	Angel Mario Lerma Sánchez, Fabiola Miroslaba Villalobos Castaldi, Adriana Mexicano Santoyo, Pedro Damián Reyes Grado de contribución igual
Metodología	Angel Mario Lerma Sánchez, Adriana Mexicano Santoyo, Fabiola Miroslaba Villalobos Castaldi, Pedro Damián Reyes Grado de contribución igual
Software	Angel Mario Lerma Sánchez (Principal), José Ruiz Pinales (Apoyo)
Validación	María Belén Nieto Castro (Principal), Angel Mario Lerma Sánchez (Apoyo)
Análisis Formal	Angel Mario Lerma Sánchez, Adriana Mexicano Santoyo, Fabiola Miroslaba Villalobos Castaldi, Pedro Damián Reyes Grado de contribución igual
Investigación	Angel Mario Lerma Sánchez, Fabiola Miroslaba Villalobos Castaldi Grado de contribución igual
Recursos	Rogelio Cruz-Martínez (Principal),
Curación de datos	María Belén Nieto Castro (Principal), Angel Mario Lerma Sánchez (Apoyo)
Escritura - Preparación del borrador original	Angel Mario Lerma Sánchez (Principal)
Escritura - Revisión y edición	Adriana Mexicano Santoyo (Principal), Angel Mario Lerma Sánchez (Apoyo)
Visualización	Fabiola Miroslaba Villalobos Castaldi (Principal), Angel Mario Lerma Sánchez (Apoyo)
Supervisión	Adriana Mexicano Santoyo (Principal), Angel Mario Lerma Sánchez (Apoyo)
Administración de Proyectos	Angel Mario Lerma Sánchez, Adriana Mexicano Santoyo, Fabiola Miroslaba Villalobos Castaldi, Pedro Damián Reyes Grado de contribución igual
Adquisición de fondos	Angel Mario Lerma Sánchez