# Classificando Graus de Pterígio Utilizando Aprendizado de Máquina

# Classifying Degrees of Pterygium Using Machine Learning

**Clasificación de Grados de Pterigión Mediante Aprendizaje Automático**

Márcia Ito1, Lincoln de Assis Moura Jr.2

1 Professora pleno do Departamento de Tecnologia e Informação, Faculdade de Tecnologia de São Paulo – Fatec-SP, São Paulo (SP), Brasil.

2 Diretor, vertical Saúde da Accenture do Brasil, São Paulo (SP), Brasil

Autor correspondente: Márcia Ito

E-mail: marcia.ito01@fatec.sp.gov.br

## Resumo

Técnicas de inteligência computacional (IC) têm sido amplamente utilizadas em diagnóstico médico assistido por computador ao longo dos últimos anos, com destaque especial às Redes Neurais Artificiais (RNAs). Neste trabalho, uma abordagem utilizando Rede Neural Convolucional (CNN) para a classificação de graus de avanço de pterígio, para o auxílio de diagnóstico médico, é apresentada. A rede CNN possui 5 camadas de extração de características e 2 camadas totalmente conectadas para classificação. A CNN proposta obteve precisão, *recall* e F1-*score* igual a 91%, 91% e 91% respectivamente. Até o desenvolvimento deste trabalho, este é o primeiro resultado de alta precisão obtido na literatura para o mesmo propósito.

**Descritores:** Pterígio; Patologia; Olho; Rede Neural Convolucional; Classificação

## Abstract

Computational intelligence (CI) techniques have been widely used in computer-assisted medical diagnosis over the past few years, with special emphasis on Artificial Neural Networks (ANNs). In this work, an approach using Convolutional Neural Network (CNN) for the classification of degrees of pterygium advancement, for the aid of medical diagnosis, is presented. The CNN network has 5 layers of feature extraction and 2 layers fully connected for classification. The proposed CNN achieved accuracy, recall and F1-score equal to 91%, 91% and 91% respectively. Until the development of this work, this is the first high precision result obtained in the literature for the same purpose.

**Keywords:** Pterygium; Pathology; Eye; Convolutional Neural Network; Classification

## Resumen

Las técnicas de inteligencia computacional (IC) se han utilizado ampliamente en el diagnóstico médico asistido por computadora durante los últimos años, con especial énfasis en las redes neuronales artificiales (ANN). En este trabajo, se presenta un enfoque que utiliza la red neuronal convolucional (CNN) para la clasificación de los grados de avance del pterigión, para la ayuda del diagnóstico médico. La red de CNN tiene 5 capas de extracción de características y 2 capas completamente conectadas para la clasificación. La CNN propuesta logró precisión, recuperación y puntuación F1 iguales al 91%, 91% y 91%, respectivamente. Hasta el desarrollo de este trabajo, este es el primer resultado de alta precisión obtenido en la literatura con el mismo propósito.

**Descriptores:** Pterigión; Patología; Ojo; Red neuronal convolucional; Clasificación

## Introdução

O pterígio é um crescimento fibrovascular resultante de tecido subconjuntival estendendo-se através do limbo até a córnea. A taxa de prevalência de pterígio primária varia de 0,7 a 30% em várias populações ao redor do mundo, onde a maior porcentagem dos casos ocorrem em regiões equatoriais. (1) Dentre as causas do surgimento dessa patologia, a principal é a exposição prolongada aos raios Ultra Violeta (UV). (2) O pterígio possui 4 graus que são classificados com base no tamanho do avanço do tecido fibrovascular ao longo do olho, sendo esses graus definidos como: Grau 1 (menor que 2 mm); Grau 2 (entre 2 a 4 mm); Grau 3 (entre 4 a 6 mm); Grau 4 (maior que 6 mm). (3) A Figura 1 mostra os graus de avanço em olhos afetados.

Pterígio é uma doença ocular prevalente entre os trabalhadores que são frequentemente expostos a luz solar como pescadores, fazendeiros e construtores civis. A maior parte dessas pessoas raramente estão cientes da existência do pterígio. Portanto, profissionais da saúde geralmente as visitam para realizar exames básicos de saúde que incluem o rastreamento de doenças oculares. Devido a muitos testes que precisam ser realizados para cada ciclo de verificação, uma ferramenta moderna é necessário para tornar o processo de triagem mais rápido. Além disso, ela deve também apresentar um diagnóstico primário mais completo, onde deve informar tanto se existe um pterígio no olho do paciente, como também discriminar o grau da patologia.



**Figura 1 –** Exemplos de Graus de avanço de pterígio (4)

No intuito de auxiliar médicos no diagnóstico de doenças oculares, como pterígio, são desenvolvidos Sistemas de Auxílio ao Diagnóstico (*Computer-Aided Diagnosis* - CAD). Usualmente, esses sistemas utilizam técnicas de processamento de imagens e aprendizado de máquina para classificar imagens como saudáveis ou doentes, provendo assim um diagnóstico suplementar. Eles também podem auxiliar no atendimento de mutirões e na prevenção de fadiga por parte do especialista, o que pode evitar a ocorrência de diagnósticos errôneos. (5,6) Para este trabalho em particular, tem-se interesse em utilizar uma Rede Neural Convolucional (CNN) para a classificação de graus de avanço de pterígio em olhos afetados por essa patologia.

Apesar do grande uso de CNN como um dos principais meios de diagnósticos automáticos de patologias oftalmológicas, de 2015 até a data de implementação deste projeto, verificou-se que sua utilização para a classificação desta patologia em específico não é amplamente estudado. Foram encontrados na literatura somente 4 artigos que tratam da classificação desta patologia, no entanto eles apenas classificam as imagens em pterígio e não pterígio (7,8,9), ou apresentam métodos para segmentar e identificar a doença (10). A literatura não possui qualquer trabalho que apresente um método completo sobre a classificação dos graus desta doença.

Uma possível justificativa para esse número pequeno de artigos na classificação de pterígio, e carência na classificação dos graus de avanço dessa patologia, é a escassez de uma base de dados de fácil acesso com amostras em imagens de olhos afetados por esta doença que permitam a formalização do problema e treinamento de uma RNA.

Diante disso, o presente artigo descreve a implementação de uma Rede Neural Convolucional (CNN) para a classificação de graus de avanço de pterígio. A metodologia adotada baseia-se em: obtenção de uma base de dados que possui imagens de olhos saudáveis e afetados com pterígio em seus mais diversos graus de avanço; pré-processamento dessas imagens com uma técnica de redimensionamento e *data augmentation*; modelagem e treinamento de RNAs; comparação entre as redes desenvolvidas; testes; e documentação dos resultados.

Além disso, vale destacar que o presente artigo contribui não só com um modelo CNN original para classificação de graus de pterígio, como também a disponibilização de uma base de dados já pré-processada para treinamento de futuros novos modelos.

**Métodos**

Nesta seção, são apresentados os materiais e o método proposto neste trabalho. O método consiste basicamente de quatro etapas: obtenção da base de imagens, pré-processamento das imagens, classificação das imagens, testes e comparação dos resultados. A Figura 2 ilustra o fluxo geral do método e nas próximas subseções cada uma dessas etapas será apresentada detalhadamente.

**Figura 2 –** Fluxograma do projeto

**Base de Dados Pterígio**

Para projetar classificadores automáticos capazes de discriminar vários graus de pterígio, faz-se necessário definir um conjunto de dados tanto com imagens de olhos saudáveis, bem como de acometidos por essa patologia nos mais variados graus de avanço. Assim, como uma estrutura de dados que possua a classe que cada imagem pertence (grau 0, grau 1, grau 2, grau 3 ou grau 4). Este *dataset* pode então ser usado para treinar e validar diversos classificadores.

O conjunto de dados utilizado neste é composto por 390 imagens com extensão JPEG (*Joint Photographic Experts Group*). Cada imagem possui as dimensões de 3000 x 3000 *pixels* RGB (*Red Green Blue*). As imagens foram disponibilizadas pela clínica oftalmológica Maranguape (11) e capturadas utilizando um auto refrator *greens* TCI. A base de dados possui 28 amostras para olhos sem a patologia (Grau 0); 20 amostras para pterígio tipo 1; 40 para tipo 2; 202 para tipo 3; 100 para tipo 4. Além disso, cada imagem foi categorizada por especialistas na área.

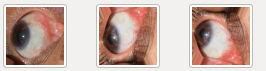
## Pré-processamento dos Dados

O pré-processamento das imagens é realizado em duas subetapas. A primeira subetapa consiste no redimensionamento das imagens para o tamanho 128x128 com o objetivo de reduzir o custo computacional do método. Ressalta-se que foram testados outros tamanhos de imagens, porém, o tamanho 128x128 foi o que obteve a melhor relação de custo e desempenho.

Para a tarefa de classificação de imagens, redes CNNs são treinadas utilizando volumes gigantescos de amostras de entradas, milhares ou até milhões. Além disso, cada classe dessa base de dados deve possuir amostras nos mais variados ângulos, *zoom* e qualidade. Isso é feito principalmente para se evitar o *overfitting*, que ocorre quando o modelo se ajusta demais ao treinamento, mas não tem bom desempenho na etapa de validação.

A base de dados inicial apresentada neste trabalho carece drasticamente tanto de volume de dados, quanto de heterogeneidade das imagens. Assim, uma abordagem alternativa para contornar isso foi, após o redimensionamento, processar as imagens utilizando uma técnica conhecida como *data augmentation*. Nessa abordagem, cada imagem passa por um processo de “clonagem”, no entanto, os clones recebem tratamentos para ficarem ligeiramente diferentes das imagens originais. Para cada imagem foram geradas novos exemplares rotacionados, transladados, expandidas ou contraídas, e com variações na cor e ruídos. Assim, um *dataset* mais heterogêneo pôde ser obtido.

Para este trabalho, foram testados várias quantidades diferentes de exemplares gerados no *data augmentation* com intuito de analisar qual o valor mínimo necessário para se ter um modelo confiável e com bom desempenho. Assim, o valor escolhido foi um de 25x (foram geradas 25 amostras para cada imagem da base de dados), logo, a base de dados final tinha 9750 imagens. A Figura 3 mostra um exemplo do resultado do pré-processamento de uma imagem.



**Figura 3 –** Exemplo de um imagem redimensionada e processada com o *data augmentation*

## Classificação das imagens

A classificação das imagens foi realizada utilizando uma Rede Neural Convolucional (CNN), que são modelos biologicamente inspirados nos processamentos de dados visuais que podem aprender características de maneira hierárquica. (12)

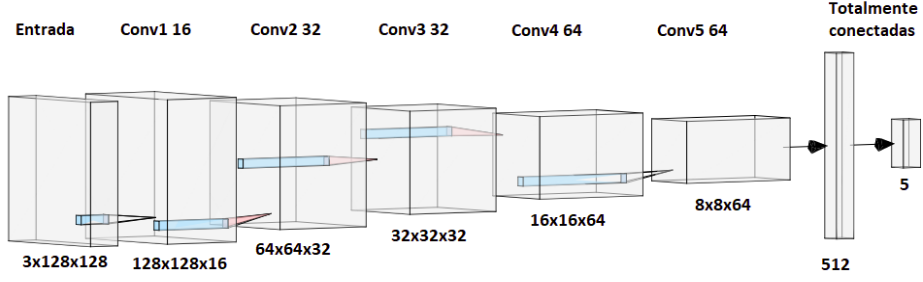
CNN´s possuem camadas convolucionais. Nelas, a saída de cada neurônio em uma camada convolucional é apenas uma função de um subconjunto (normalmente pequeno) dos neurônios da camada anterior. Isso faz as CNN´s serem mais poderosas e ao mesmo tempo terem um custo de tempo de execução muito menor se comparado a uma rede *MultiLayer Perceptron* (MLP). (13)

As camadas convolucionais são responsáveis por extrair características de cada padrão de entrada proveniente da camada anterior. Sua grande vantagem é a extração de características locais de subáreas dentro de cada amostra, essa abordagem aumenta sua capacidade de discriminação dos dados. Já as camadas totalmente conectadas, que ficam no fim da rede, são utilizadas para classificar os padrões e extrair os pesos da rede. (14)

Neste trabalho, uma CNN compacta foi desenvolvida especificamente para classificar se existe a presença ou não da patologia pterígio em imagens de olhos e também identificar qual o grau da patologia em caso de ocorrência da mesma. Foi utilizada a biblioteca *keras tensor flow* disponível na linguagem de programação *python* 3.7 para sua modelagem. O modelo da rede foi desenvolvido como se segue:

* + Tamanho da entrada: 128x128x3.
  + Camada convolucional 1: Filtro 3x3; 16 neurônios; função de ativação ReLu; *max pool* 2x2; *dropout* 20%.
  + Camada convolucional 2: Filtro 3x3; 32 neurônios; função de ativação ReLu; *max pool* 2x2; *dropout* 25%.
  + Camada convolucional 3: Filtro 3x3; 32 neurônios; função de ativação ReLu; *max pool* 2x2; *dropout* 25%.
  + Camada convolucional 4: Filtro 3x3; 64 neurônios; função de ativação ReLu; *max pool* 2x2; *dropout* 25%.
  + Camada convolucional 5: Filtro 3x3; 64 neurônios; função de ativação ReLu; *max pool* 2x2; *dropout* 25%.
  + Camada dense 1: 512 neurônios; função de ativação *sigmoid*; *dropout* 50%.
  + Camada dense 2: 5 neurônios; função de ativação *softmax*.

O porcentagem de imagens selecionadas para treino e teste foram, respectivamente, 80% e 20%, onde desses 80% de amostras para treino, 20% foram usadas para validação do modelo. A seleção das imagens foi feita de maneira aleatória. O modelo foi treinado por 60 épocas, onde foi utilizado um tamanho de *batch* de 64, função de perda entropia cruzada e otimizador Adam. A Figura 4 ilustra a arquitetura da CNN desenvolvida.

**Figura 4 –** Arquitetura CNN utilizada neste projeto

Outros dois classificadores também foram desenvolvidos com intuito de comparar seus resultados na classificação das imagens e obter maior confiança quanto a escolha do melhor modelo de aprendizado de máquina para este problema. entre os métodos. O Primeiro é baseado em um *MultiLayear Perceptron* (MLP). Ele foi arquitetado com 5 camadas intermediárias, cada uma tendo 8 neurônios, e foi treinado por 60 épocas. O segundo foi baseado em um classificador logístico multi classe. Neste, o *solver* escolhido foi o *Stochastic Average Gradient* (SAG) (aconselhado para grandes volumes de dados) e o classificador foi treinado por 1000 épocas.

## Testes

Para ambos os classificadores, a base de dados foi dividida em 80% para treino e validação, enquanto os outros 20% foram utilizados para testes. Vale ressaltar que a base de dados foi previamente embaralhada antes da etapa de divisão da mesma.

## Comparação

Para comparar os classificadores desenvolvidos, as métricas comumente utilizadas e aceitas pela comunidade científica para análise de desempenho de sistemas CAD são a especificidade, a precisão (*precision*), sensibilidade (*recall*) e F1-*score*. (16) Lembrando que essas métricas são calculadas a partir da contagem de verdadeiros positivos, falsos positivos, falsos negativos e verdadeiros negativos. Além disso, como este trabalho engloba um problema de classificação multiclasse, a *weighted average* foi escolhida como média das métricas de avaliação das classes. Além disso, a métrica especificidade não será explicitamente citada, visto que para para casos onde o olho não possui pterígio, o classificador apenas deve classificar como pertencente a classe 0. Assim, a especificidade será baseada nos resultados dessa classe.

## Resultados e Discussão

Como detalhado na Seção de Métodos, a discriminação de pterígio, e seus graus, é obtida a partir do treinamento de um classificador que tenha suporte a lidar com problemas que envolvem imagens e várias classes. Os resultados mais promissores, como previsto, foram os obtidos pela CNN desenvolvida. Na etapa de teste foram obtidos 91% de sensibilidade (*recall)*, 91% de *precision* e 91% de F1 *score*.Os resultados obtidos destes testes estão detalhados na Tabela 1. Pode ser notado que, de fato, aumentar a base de dados em 25 vezes permite um aumento significativo na taxa de sensibilidade (*recall*) do modelo CNN.

**Tabela 1 –** Comparação entre o*s* resultados dos classificadores

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Classificador** | **épocas** | ***recall*** | ***precision*** | **F1 s*core*** |
| **CNN** | 60 | 0,96 | 0,96 | 0,96 |
| **MLP** | 60 | 0,47 | 0,22 | 0,31 |
| **Logístico** | 1000 | 0,58 | 0,57 | 0,57 |

A Tabela 2 exibe o resultado da classificação da etapa de testes de cada classe para os classificadores desenvolvidos.*.*

**Tabela 2 *–*** F1 *score* do teste para cada classe

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Classificador** | **Classe0** | **Classe1** | **Classe2** | **Classe3** | **Classe4** |
| **CNN** | 0,99 | 0,96 | 0,96 | 0,96 | 0,94 |
| **MLP** | 0,00 | 0,00 | 0,00 | 0,64 | 0,00 |
| **Logístico** | 1,00 | 0,35 | 0,32 | 0,67 | 0,42 |

## Como pode ser observado na Tabela 1 e 2, o uso de CNN para este trabalho é puramente justificável tanto pela sua grande capacidade de generalização dos dados amostrais, quanto pela capacidade de classificar de forma confiável cada classe se comparado com outros métodos clássicos de discriminação disponíveis. Além disso, no Quadro 1 pode ser observado algumas comparações realizadas com outros trabalhos encontrados na literatura.

**Quadro 1 –** comparação entre métodos encontrados na literatura

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Autores e ano** | **Classificador** | **Tamanho**  ***data set*** | ***Data***  ***augmentation*** | **Taxa de acerto médio** | **Treino** | **Classes** |
| Método proposto | CNN | 390 | Sim (25x) | 96% | 80% | 5 |
| Zaki et al. (2019) | SVM | 3017 | Não | 91,7% | 80% | 2 |
| Zulkifley et al. (2019) | CNN | 120 | Não | 98,3% | 80% | 2 |
| Pagnin (2011) | OPF | 7654 | Não | 99,41% | 75% | 2 |

Com base nos resultados observados no Quadro 1, pode ser observado que os trabalhos de Pagnin (8) (que utiliza Floresta de Caminhos Ótimos em seu trabalho) e Zulkifley et al (9) (que utiliza máquina de vetores de suporte) obtiveram uma taxa de acerto médio maior que o método aqui apresentado, no entanto vale ressaltar que ambos classificam apenas os olhos em acometidos ou não com pterígio. Isso faz com que a complexidade do problema deles se torne linear, logo de mais fácil solução. Apesar dessas observações, a diferença entre a taxa de acerto médio deste trabalho e a de Pagnin (8) é de somente 4,41%. Assim pode ser inferido que o método aqui apresentado possui uma maior robustez dado ao número maior de classes a serem discriminadas.

Vale ressaltar que tal comparação com a literatura se torna bastante difícil, não só pelo fato de ser categorias de metodologias diferentes, mas também pelo número e pela base de imagens utilizada nos trabalhos, que nem sempre são públicas. Apesar dos esforços em testar a metodologia proposta nas demais bases, isto não foi possível, visto há dificuldade da aquisição das bases privadas.

## Conclusão

Este artigo apresentou um método para classificação de pterígio em imagens de olhos sem e com esta patologia, além de classificar o grau da doença. Sendo baseado em uma CNN, o método desenvolvido utiliza uma rede compacta e *data augmentation* para suprir a falta de amostras da base de dados.

Dado que a classificação automática de patologias oculares, no caso deste trabalho a de pterígio, fornece uma ferramenta a mais para apresentar um diagnóstico mais rápido e com uma taxa de erro humano bem menor, o trabalho desenvolvido teve como objetivo fornecer um método robusto para a classificação de graus de avanço de pterígio a partir de imagens oculares. Desta forma, a base de dados utilizada e o método desenvolvido poderão ser utilizados em processos de diagnóstico médico assistido por imagem. Assim, o método poderá fornecer auxílio a especialistas no processo de triagem e diagnóstico desta patologia.

O método desenvolvido obteve resultados promissores, alcançando uma sensibilidade nos testes de 91% com *data augmentation* de 25x, e valores de perda nos testes de 13,68%. Entretanto, o método ainda poderá ser melhorado. Como possíveis trabalhos futuros, é pretendido investir em uma nova base de dados contendo um volume maior de dados. Além disso, técnicas de pré-processamento poderão ser utilizadas para melhorar as imagens nas etapas iniciais. Finalmente, o método proposto também deverá ser validado em outras bases, com o objetivo de verificar sua capacidade de generalização.

## Agradecimentos

Agradecimentos aos membros do comitê científico do CBIS 2020 que auxiliaram no desenvolvimento deste *template*.

À instituição SBIS, pela realização das medidas ou empréstimo de equipamentos.

À FAPESP (ou CNPq ou outra), pelo apoio financeiro ou por bolsas concedidas.

Ao Professor/Pesquisador/Técnico XXX pela colaboração.

## Referências

1. Wazlawick R. Metodologia de Pesquisa para Ciência da Computação. 2a. ed. São Paulo: GEN LTC; 2017. 150 p.
2. Pereira MG. Dez passos para produzir artigo científico de sucesso [Internet]. Epidemiol. Serv. Saude 2017 jul./set.;26(3) doi: http://dx.doi.org/10.5123/s1679-49742017000300023[cited 2020 Jul 28]. Available from: <http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742017000300661>
3. Caramelli B. Conclusão: como exibir a cereja do bolo [Internet]. Ver. Assoc. Med. Bras. 2012 nov./dec.; 58(6) doi: http://dx.doi.org/10.1590/S0104-42302012000600001 [cited 2020 Jul 28]. Available from: <https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-42302012000600001>
4. Aquino IS. Como escrever artigos científicos: sem Arrodeio e sem medo da ABNT. 8a. ed. São Paulo:Saraiva; 2012. 126 p.
5. Ito M, Moyano LG, Appel AP, de Santana VF. Análise do relacionamento na comunidade de médicos de seguradoras de saúde. Anais do XV Congresso Brasileiro de Informática em Saúde. 2016 Nov 27-30; Goiania. 309-18.
6. Campos Filho AS, Lemos WB, Souza, RC, Lima, LLB. Realidade virtual como ferramenta educacional e assistencial na saúde: uma revisão integrativa.J.Health Inform. 2020 Abril-Junho; 12(2):58-63.
7. Da Silva MA, Santos N, Rosa COCS, Santos M, Ito M, Vieira A, Ziviani A, Oliveira RM. Análise dos atendimentos de gestantes na rede de atenção básica de saúde no município de São Paulo. Anais do XX Simpósio Brasileiro de Computação Aplicada a Saúde. 2020 Sept 15-18; Salvador. (no prelo)
8. Ito C, Haddad AE, Pitta Neto L., Ito M. OdontoGame: Um jogo sério para conscientização de higiene bucal de crianças em idade pré-escolar. Proceedings of XVIII SBGames - Game and Health Workshop.2019 Oct 28; Rio de Janeiro. 1408-9.
9. Saraiva EH, Patrao DFC, Ito M. Analysis of usage of term weighting algorithm for mapping health procedures into the unified terminology of supplemental health (TUSS). Stud Health Technol Inform. Medinfo 2019: Health and Wellbeing e-Networks for All. 2019: 1496-7. doi: 10.3233/SHTI190502