

Aplicação de imagens sintéticas para otimização de modelos computacionais de detecção do estrabismo

Jonathan S. Santos¹, Ismar Silveira Frango¹

¹Universidade Presbiteriana Mackenzie (UPM)
R. da Consolação, 930 – 01302-907 – São Paulo – SP – Brasil

silva.santos.jonathan@gmail.com, ismarfrango@gmail.com

Abstract. *Strabismus is among the eye diseases that most lead to blindness or low vision, affecting about 4 % of the world population. Fortunately, the disease can be treated. Diagnosis even in the first moments of its manifestation dramatically increases the possibility of successful treatment. There are several proposals in the scientific literature for detecting and supporting the diagnosis of pathology, however, we have not found studies that seek to propose means of optimization for these techniques. This article presents a methodology for optimizing supervised strabismus detection models by increasing data using realistic synthetic samples. In evaluation, the proposed technique resulted in a gain of 7 % accuracy.*

Resumo. *O Estrabismo está entre as doenças oculares que mais levam à cegueira ou baixa visão, afetando cerca de 4% da população mundial. Felizmente a doença pode ser tratada. Diagnósticos ainda nos primeiros momentos de sua manifestação aumentam drasticamente a possibilidade de sucesso do tratamento. Há diversas propostas na literatura científica para a detecção e apoio ao diagnóstico da patologia, no entanto não encontramos trabalhos que buscam propor meios de otimização para essas técnicas. Este artigo apresenta uma metodologia de otimização de modelos supervisionados de detecção de estrabismo através do aumento de dados utilizando amostras sintéticas realistas. Em avaliação, a técnica proposta resultou em um ganho de 7% de acurácia.*

1. Introdução

1.1. Objetivo

O principal objetivo desta pesquisa é contribuir com os métodos e aplicações computacionais existentes para a detecção do *estrabismo*, através da proposta de um modelo de aprendizagem profunda, generativo, capaz de gerar imagens sintéticas da região dos olhos de pessoas com *estrabismo*, em função da oportunidade de uso dessas imagens para a otimização de técnicas de aprendizagem de máquina supervisionadas.

1.2. Justificativa

A publicação de trabalhos que buscam trazer tecnologias da computação para o *estrabismo* ainda são numericamente diminutas. A principal razão que motiva tal escassez deve-se ao fato de não haver base de dados de *estrabismo* pública, portanto, os trabalhos existentes utilizam bases privadas [Lu et al. 2018, p.2].

Este artigo além de ter como justificativa a escassez de base de dados públicas, também justifica-se diante da compreensão de haver uma oportunidade de aumentar base já existentes, por meio do emprego de técnicas de *data augmentation*.

A pesquisa parte da hipótese - comprovada de forma experimental - que o *aumento de dados* com amostras sintéticas pode aumentar a performance de técnicas de detecção do estrabismo.

1.3. Motivação

Motiva os autores desta pesquisa a oportunidade de contribuir com demais pesquisas científicas, entendendo estes existir condições para novas propostas técnicas que podem ter efeito prático e benéfico na vida de pacientes estrábicos. Também são motivantes os desafios técnicos que se apresentam no horizonte projetado pelas pesquisas precursoras na área, assim como as inovações emergentes em computação.

1.4. Visão geral

É proposto nesta pesquisa um modelo de geração de imagens sintéticas fotorrealistas da região dos olhos de pessoas com Estrabismo.

É explorado também neste estudo o potencial ganho de performance em modelos de aprendizagem de máquina supervisionada, quando aplicadas técnicas de aumento de dados (*data augmentation*) com a adição de imagens sintéticas em bases de dados pré-existentes.

Os aspectos de geração de amostras sintéticas foram já apresentados com maiores detalhes em um artigo dedicado a estes fins denominado *Generating photorealistic images of people's eyes with strabismus using Deep Convolutional Generative Adversarial Networks* [Santos and Silveira 2020].

2. Estrabismo

2.1. Introdução ao Estrabismo

Estrabismo é a disfunção em que ocorre o desalinhamento dos olhos, ocorre quando os eixos dos olhos não direcionam-se para o mesmo ponto. O Estrabismo é uma patologia geralmente diagnosticada durante a infância. A primeira suspeita do doença surge comumente com os pais ou com o pediatra, sendo o profissional em oftalmologia a autoridade mais reconhecida e indicada para gerar um diagnóstico preciso e prescrever o tratamento [Lopes 2006].

2.2. Causas

As causas do estrabismo relacionam-se a maneira como os olhos fazem suas operações motrizes ou, em outras palavras, a forma com que os olhos se movimentam. Seis músculos em cada olho são responsáveis pela movimentação ocular. A dessincronia, ou desajuste entre esses músculos em relação aos olhos são as causas comuns de sua origem. A natureza que implica o fator de desajuste pode ser de fundamento físico, onde os músculos podem estar atrofiados ou lesionados, ou de caráter neurológico. O controle dos músculos oculares é feito através de pulsos nervosos emitidos pelo cérebro, portanto em quadros de

patologias que afetam o cérebro como: hidrocefalia, síndrome de down, cânceres intracranianos e viroses. As principais causas que levam ao quadro, na infância, são: síndromes de Noonan Apert, Alport, trissomia do 18, Down, Prader-Willi, Duane, incontínência pigmentar; Além de infecções como rubéola congênita, toxoplasmose [Lopes 2006]. Já na fase adulta, temos: Baixa visão, diabetes, trombose de seio cavernoso, paralisias de nervos cranianos, fraturas de órbita, Guillain-Barré, botulismo, envenenamentos, lesões cerebrais, tumores, hidrocefalia, aneurismas, miastenia gravis, doença de Graves, traumas, fraturas de órbita [Lopes 2006].

2.3. Tratamento

O diagnóstico e tratamento do estrabismo na infância podem ser primordiais para o desenvolvimento visual saudável. O desenvolvimento visual normal durante a infância pode ser inibido por grandes erros de refração, anisometropia e estrabismo, e as conseqüências prejudiciais para a visão podem ser duradouras [Ingram 1977]. O tratamento do estrabismo, quando identificado na infância, deve ser iniciado o mais precocemente possível para se evitar a ambliopia, podendo envolver a utilização de óculos, oclusores ou mesmo cirurgia. Não se deve aguardar para encaminhar a criança ao oftalmologista; desvios observados após o terceiro mês de vida devem ser avaliados.

Em adultos, duas situações diferentes podem ocorrer: o estrabismo presente desde a infância, com tratamento semelhante ao discutido anteriormente ou aquele adquirido depois do sistema visual desenvolvido. Nestes casos, a ocorrência de diplopia (visão dupla) é freqüente e geralmente incapacitante. Podem advir principalmente de perda da visão de um olho, de lesões musculares, de lesões nervosas ou vasculares. O tratamento depende do tipo e da amplitude do desvio e pode envolver oclusores, cirurgia, óculos e prismas, na maioria dos casos. Os desvios são divididos em: desvios latentes (forias) ou manifestos (tropias); desvios constantes ou intermitentes; desvios dos olhos divergentes (exotropias), convergentes (esotropias) ou desvios verticais. Todos os tipos de estrabismo podem levar a disfunção visual e necessitam de atenção quando secundários a qualquer patologia. O tratamento do estrabismo não deve ser considerado apenas do ponto de vista estético, em especial em adultos, pois além do combate à ambliopia ou à diplopia, há ganho na coordenação dos olhos, com melhora do campo de visão binocular, na percepção de profundidade, além dos ganhos óbvios em auto-estima, oportunidades de trabalho e socialização [Lopes 2006].

3. Fundamentação Teórica

3.1. Redes Neurais Convolucionais

As redes neurais convolucionais foram originalmente projetadas para solucionar o desafio de reconhecer dígitos manuscritos. [Fukushima 1980, LeCun et al. 1989]

Trata-se de uma arquitetura de rede neural artificial onde há processos convolucionais conectados às camadas da rede. Os processos convolucionais podem ocorrer em N dimensões e podem ser compreendidos como sendo multiplicações de matrizes. Os valores de entrada do modelo formam a matriz de N dimensões que será multiplicada por N matrizes denominadas *kernels* ou *filtros*. Os filtros utilizados em redes convolucionais são comumente análogos àqueles relacionados a processos de computação gráfica como

realce de contornos, aumento de contrastes, entre outros. Os *filtros* são responsáveis por discretizar diferentes aspectos de uma entrada.

Realizar um processo convolucional em uma imagem, por exemplo, é equivalente a passar um detector de características visuais (filtro) local pela imagem e registrar todos os locais na imagem em que a característica visual estava presente. A saída desse processo é um mapa de todos os locais na imagem em que a característica relevante ocorreu. Por esse motivo, a saída de um processo de convolução é também conhecida como mapa de características. [Kelleher 2019, p.159]

As características discretizadas pelos *filtros* se tornam os valores utilizados como entrada para a camada de uma rede neural artificial. [Goodfellow et al. 2016, p.322].

3.2. Redes Adversárias Generativas (GAN)

As *redes adversárias generativas (GAN)*, propostas em [Goodfellow et al. 2014] são fundamentadas em uma arquitetura com o propósito de construir modelos generativos, que, por sua vez, são modelos que tentam aprender a distribuição dos dados (ou, modelar o processo que gerou os dados). [Kelleher 2019, p.234].

As *GAN* são constituídas por uma estrutura que contém em duas redes neurais que aprendem juntas por meio da competição mútua. Existe a rede *D* que cumpre o papel do *discriminador*, responsável por identificar se um objeto em particular faz parte de uma classe específica: um classificador binário. E existe também a rede *G*, que é o *gerador*, responsável por fornecer uma entrada feita de valores aleatórios que formam um objeto que será então avaliado pela rede *D*. [Goodfellow et al. 2014, p.1].

A dinâmica entre redes é a mesma que em um jogo *Minimax* de dois jogadores, onde a rede geradora procura gerar objetos cada vez mais difíceis de distinguir pela rede discriminadora. Com dificuldade crescente, a rede discriminadora aprimora cada iteração de objetos de classificação, decidindo se eles pertencem ou não a uma classe específica. [Goodfellow et al. 2014, p.1].

O modelo gerador pode ser pensado como análogo a uma equipe de falsificadores, tentando produzir moeda falsa e usá-la sem serem descobertos, enquanto o modelo discriminador é análogo à polícia, tentando detectar a moeda falsa. A competição neste jogo impulsiona ambas as equipes aprimorarem seus métodos até que as falsificações sejam indistinguíveis das originais. [Goodfellow et al. 2014, p.1].

As Redes Adversárias Generativas Convolucionais Profundas (DCGAN) são um uso prolongado do GAN [Radford et al. 2015], construindo suas redes neurais em uma arquitetura de aprendizado profundo; mais especificamente: redes convolucionais.

4. Trabalhos Correlatos

Nesta seção está presente uma seleção de trabalhos acadêmicos que de algum modo se relacionam com este artigo, sejam essas relações com estrabismo ou relações com uso de amostras sintéticas como um artifício de *data augmentation*.

4.1. Photorefraction: a useful tool to detect small angle strabismus

Em [Abrahamsson et al. 1986], Abrahamsson et al., usam fotorrefração para obter a detecção de estrabismo de ângulo pequeno. Quarenta pacientes com estrabismo de ângulo

pequeno e 30 controles de olhos retos foram examinados por foto-refração, a fim de estabelecer a validade desse método para detectar o estrabismo. A sensibilidade do método foi baseada na detecção de assimetria da luz da córnea e / ou reflexos do fundo. Todos os casos com desvio de 2 graus ou mais foram detectados e 2 controles foram falsamente considerados pacientes estrábicos. A sensibilidade do método diminuiu quando o paciente usou a correção ideal. Utiliza o método para detecção de refração com fotografias desenvolvido por [Kaakinen 1979]. Trata-se de um *short paper* que explora o uso da foto-refração para identificação e triagem de estrabismo em crianças. Este estudo não envolve um uso computacional, as análises são feitas através de fotografias analógicas e testes humanos. A conclusão deste estudo é que todos, os pacientes com um pequeno desvio angular de dois graus ou mais poderia ser corretamente identificados a partir das fotografias. A assimetria nos reflexos do fundo e da córnea ainda era detectável quando o paciente usava a correção ideal. As lentes esféricas e cilíndricas diminuíram a sensibilidade do método para detectar o estrabismo. Os pacientes que tiveram correção com lentes prismáticas não tiveram estrabismo reconhecido uma vez que foi impossível de identificar os reflexos do feixe de luz nas fotografias.

4.2. Computational methodology for automatic detection of strabismus in digital images through Hirschberg test

Em [De Almeida et al. 2012], De Almeida et al. aplicam uma câmera digital e o teste de Hirschberg para identificar estrabismo. Nesta pesquisa Apresentam uma metodologia para detecção automática de estrabismo em imagens digitais através do teste de Hirschberg. O trabalho foi organizado em quatro etapas: (1) encontrar a região dos olhos; (2) determinar a localização precisa dos olhos; (3) localizar o limbo e o brilho; e (4) identificar estrabismo. A metodologia produziu resultados na faixa de 100% de sensibilidade, 91,3% de especificidade e 94% para a correta identificação do estrabismo. Esta pesquisa foi realizada com o apoio do Dr. Jorge Meireles, especialista em estrabismo. Um dos fundamentos do método proposto é o desenvolvimento de uma solução barata para o diagnóstico do estrabismo e de fácil utilização, permitindo ser operada por profissionais não especializados em estrabismo. Para a aplicação do método, bastaria ter uma máquina fotográfica digital e um computador (portátil ou não). Para o reconhecimento da região dos olhos, o artigo apresenta uma aplicação da análise de momentos de Zernike com uma máquina de suporte de vetores (SVM). A magnitude dos momentos de Zernike é usada como um recurso invariante no sentido de tornar a detecção robusta mesma para imagens rotacionadas [Kim and Kim 2008]. Momentos de Zernike são números complexos pelos quais uma imagem é mapeada em um conjunto de polinômios Zernike complexos bidimensionais [Kim and Kim 2008, Hwang and Kim 2006]. Os momentos de Zernike são codificados por um algoritmo de aprendizagem de máquina que tem a finalidade de aprender o que são padrões oculares e o que não os são.

4.3. Automatic diagnosis of strabismus in digital videos through cover test

Em [Valente et al. 2017], Valente et al. realizam detecção de estrabismo em vídeos digitais através do *cover test*. Apresentam uma metodologia computacional para diagnosticar automaticamente o estrabismo através de vídeos digitais, apresentando um teste de cobertura usando apenas um computador desktop para processar esses vídeos. O método proposto foi validado em pacientes com exotropia e consiste em oito etapas: (1) aquisição, (2) detecção da região ao redor dos olhos, (3) identificação da localização da pupila,

(4) identificação da localização do limbo, (5) rastreamento dos movimentos oculares, (6) detecção do oclutor, (7) identificação de evidências da presença de estrabismo e (8) diagnóstico. A respeito dos resultados obtidos, para detectar a presença de estrabismo, o método proposto alcançou um valor de especificidade de 100% e (2) um valor de sensibilidade de 80%, com precisão de 93,33% no diagnóstico de pacientes com extropia. Esse procedimento foi reconhecido para diagnosticar o estrabismo com um valor de precisão de 87%, reconhecendo medidas inferiores a 1 delta e um erro médio na medida de desvio de 2,57 delta.

4.4. Strabismus recognition using eye-tracking data and convolutional neural networks

Em [Chen et al. 2018], Chen et al. usam um sistema de rastreamento ocular e redes neurais convolucionais para detectar estrabismo. Neste artigo, é proposto o reconhecimento do estrabismo usando dados de rastreamento ocular e redes neurais convolucionais. Em particular, um rastreador ocular é primeiro explorado para registrar os movimentos oculares de um indivíduo. Uma imagem de desvio de olhar (GaDe) é então proposta para caracterizar os dados de rastreamento ocular do sujeito de acordo com a precisão dos pontos de olhar. A imagem do GaDe é alimentada por uma rede neural convolucional - CNN - treinada em um grande banco de dados de imagens chamado ImageNet. As saídas das camadas de conexão completas da CNN são usadas como recursos da imagem GaDe para o reconhecimento do estrabismo. Um conjunto de dados contendo dados de rastreamento ocular de indivíduos estrábicos e de indivíduos não estrábicos é estabelecido para experimentos. Resultados experimentais demonstraram que os recursos naturais da imagem podem ser transferidos para representar dados de rastreamento ocular, e o estrabismo pode ser efetivamente reconhecido pelo método proposto

O método proposto utiliza aprendizagem de máquina profunda, mais precisamente (CNN). Há também o uso de uma (SVM). A metodologia pode ser descrita da seguinte forma: inicialmente é pedido ao indivíduo testado que olhe para 9 pontos pré-determinados (supra-levo, supra, supra-dextro, levo, posição primária do olhar, dextro, infra-levo, infra, infra-dextro), fazendo com que o indivíduo focalize em pontos específicos impressos em uma tela. Enquanto o paciente olha para os distintos pontos, um rastreador ocular montado de baixo da tela registra os dados capturados dos movimentos dos olhos. Com os dados capturados pelo rastreador, são gerados 3 (três) mapas de desvio do olhar (GaDe). O primeiro mapa trata-se da fixação dos pontos pelo olho esquerdo, o segundo mapa, por sua vez, trata-se dos pontos de fixação do olho direito e, por fim, o terceiro mapa é um ponto 'virtual' estimado como ponto médio dos dois olhos. Os três mapas são representados em apenas uma imagem digital, usando os canais RGB para armazenamento dessas informações, fazendo com que cada mapa permaneça armazenado e representado exclusivamente em um canal. Em seguida, esses mapas são utilizados como parâmetros de entrada de uma CNN para que então possam ser extraídos vetores de representação (em uma rede que já fora treinada previamente com imagens extraídas da base de dados ImageNet). Com os vetores de representação à disposição, por fim é utilizado uma (SVM) para fazer a separação ótima entre as classes *estrabismo* e *não estrabismo*.

4.5. Automated Strabismus Detection for Telemedicine Applications

Em [Lu et al. 2018], Lu et al. apresentam um sistema de detecção automático de estrabismo para uso em telemedicina. Neste artigo, um conjunto de dados de tele-estrabismo

é estabelecido pelos oftalmologistas. Em seguida, é proposta uma estrutura de ponta a ponta denominada RF-CNN para obter a detecção automatizada de estrabismo no conjunto de dados estabelecido de tele-estrabismo. O RF-CNN primeiro realiza a segmentação da região ocular em cada imagem individual e classifica ainda mais as regiões oculares segmentadas com redes neurais profundas. Os resultados experimentais no conjunto de dados estabelecido de estrabismo demonstram que o RF-CNN proposto pode ter um bom desempenho na detecção automatizada de estrabismo para aplicação em telemedicina. A arquitetura *RF-CNN*. Trata-se de uma arquitetura dividida em duas etapas. Inicialmente uma CNN faz a segmentação de uma imagem que contenha uma face humana extraíndo assim o (ROI) da região dos olhos, em seguida essa área selecionada da imagem é utilizada em outra CNN específica para a classificação binária.

4.6. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification

4.6.1. Visão geral

Em [Frid-Adar et al. 2018], Frid-Adar et al. usam imagens sintéticas geradas com (GAN) para aumentar a performance de *redes neurais convolucionais* que classificam imagens médicas de diferentes tipos de lesões hepáticas. Neste artigos são apresentados métodos para gerar imagens médicas sintéticas usando as Redes Adversárias Generativas - (GAN) - de aprendizado profundo, primeiramente apresentadas em 2014. Além disso, é apresentado o apontamento que as imagens médicas geradas podem ser usadas para aumento de dados sintéticos e apresenta melhora do desempenho da CNN para classificação de imagens médicas. O método é demonstrado em um conjunto de dados limitado de imagens de tomografia computadorizada (TC) de 182 lesões hepáticas (53 cistos, 64 metástases e 65 hemangiomas). Primeiramente é explorado as arquiteturas (GAN) para sintetizar regiões de interesse - (ROI) - de lesões hepáticas de alta qualidade. Em seguida, é demonstrado um novo esquema para classificação de lesões hepáticas usando CNN. Por fim, é treinada a CNN usando o aumento de dados clássico e o aumento de dados sintéticos e é comparado o desempenho. Além disso, é explorado a qualidade dos exemplos sintetizados usando visualização e avaliação de especialistas. O desempenho da classificação usando apenas aumento de dados clássico rendeu 78,6% de sensibilidade e 88,4% de especificidade. Ao adicionar o aumento de dados sintéticos, os resultados aumentaram para 85,7% de sensibilidade e 92,4% de especificidade. Os autores da pesquisa acreditam que essa abordagem para o aumento de dados sintéticos pode generalizar para outras aplicações de classificação médica e, assim, apoiar os esforços dos radiologistas para melhorar o diagnóstico.

4.7. Biomedical data augmentation using generative adversarial neural networks

Em [Calimeri et al. 2017], Calimeri et al. neste trabalho propõem uma nova aplicação de - (GAN) - para a geração automática de imagens sintéticas de ressonância magnética (MRI) de partes de camadas do cérebro humano; foram realizadas avaliações quantitativas e humanas das imagens geradas, a fim de avaliar a eficácia do método. Com o apoio do Centro de Educação Nvidia GPU da Universidade da Calábria a pesquisa utiliza-se de recursos técnicos e computacionais avançados para o desenvolvimento de modelos sintetizadores de imagens biomédicas. Imagens de ressonância magnéticas de camadas do cérebro humano são sintetizadas automaticamente com o uso de (GAN), De acordo com

o método de avaliação aplicado na pesquisa, os resultados obtidos pelo modelo generativo proposto em sua metodologia são realistas se comparado com as imagens reais. A arquitetura da rede (GAN) utilizada pelos autores utiliza-se de uma composição de camadas de redes CNN em uma estrutura de rede adversária - semelhante com a abordagem de *Deep Convolutional Generative Adversarial Networks* (DCGAN). Para a validação da qualidade das imagens geradas foi utilizado uma plataforma Web ad-hoc, onde médicos e especialistas humanos foram convidados a distinguir se uma imagem é real ou gerada artificialmente. Essa abordagem abre possibilidades de geração de dados qualitativos a respeito das amostras artificiais. O conjunto de dados consiste de 46.737 (quarenta e seis mil setecentas e trinta e sete) imagens representativas de camadas cerebrais de ressonância magnética extraídas de 77 (setenta e sete) indivíduos.

5. Metodologia

5.1. Base de Dados

Neste estudo utilizamos uma base de dados que contém 296 imagens. Se trata de uma união da base *SID v2* [De Almeida et al. 2015] combinadas com imagens da região dos olhos extraídas de fotografias encontradas na base *Eye Diseases* [Eye 2020].

As imagens foram redimensionadas para o tamanho de 112 pixels de largura. Cada imagem fica com aproximadamente 42 pixels de altura, dado que Utilizamos apenas o recorte da região dos olhos. Em nossa implementação aplicamos a técnica de *zerofill* (preenchimento de pixels cujo o valor é 0 (zero) para completar a imagem), portanto, cada amostra, quando está pronta para ser utilizada pelo modelo, possui um aspecto quadrado de 112 x 112 pixels. Cada imagem possui 3 canais de cores *RGB - True Color*, cada canal com 128 bits de comprimento. Dobramos o número de amostras da base copiando cada amostra e aplicando um espelhamento horizontal. Desta forma, a base de dados passa a conter 592 amostras.

Em virtude da avaliação do potencial aumento de performance gerado pelo uso de imagens sintéticas da região dos olhos de pessoas com *estrabismo* em modelos de aprendizagem de máquina supervisionada, necessitamos coletar uma base de dados de imagens de pessoas *não estrábicas*. fizemos uso da base *Celebrities Images* [Cel 2020]. É composta de uma coleção com cerca de 100.000 imagens de faces de celebridades de todo o mundo. Selecionamos manualmente 296 imagens de pessoas que não apresentavam - ao nosso julgamento - *estrabismo* aparente. Todas as imagens foram recortadas e redimensionadas para se adequarem à porção de imagens de pessoas com *estrabismo*.

5.2. Modelo DCGAN

Propomos um modelo baseado na arquitetura de aprendizagem profunda DCGAN *Deep Convolutional Generative Adversarial Networks* [Radford et al. 2015]. A implementação do software que contém a arquitetura fundamental para o modelo que propomos, é baseada no trabalho de [Ker 2020]. A maior parte das adaptações aplicadas diz respeito ao redimensionamento das redes neurais profundas em função de suportarem as amostras da base de dados que dispomos. Novas camadas foram adicionadas às redes *geradora* e *discriminadora*.

A Figura 1 apresenta um diagrama simplificado que ilustra as estruturas das redes neurais que fazem da metodologia apresentada.

5.3. Classificador binário

Desenvolvemos um classificador binário para classificar imagens entre duas classes: *estrabismo*, *sem estrabismo aparente*. Trata-se de um modelo para que possamos avaliar o potencial do uso de imagens sintéticas de estrabismo como fator que possa aumentar sua performance. A Figura 1 apresenta diagrama com uma representação simples de cada camada do modelo em questão.

5.4. Parâmetros dos modelos e implementação

Do mesmo modo que ocorre com as redes que formam a arquitetura *DCGAN*, o modelo classificador é sequencial, sua função de perda (loss) é *entropia cruzada* [Los 2020], o otimizador é *Adam* [Opt 2020] e a métrica é a *acurácia* [Met 2020].

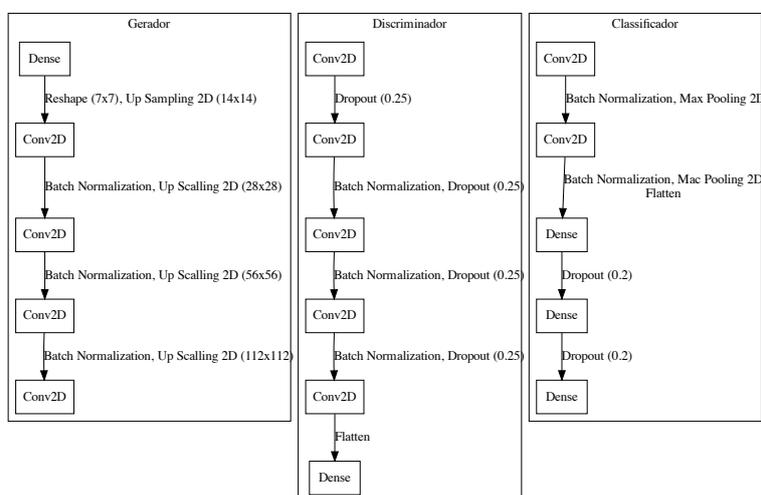


Figura 1. Representação simplificada das redes neurais geradora, discriminadora e classificadora.

A implementação, utiliza *Python* [pyt 2020] 3.6.8. Construímos os modelos de aprendizagem profunda utilizando o framework *Keras* [Hom 2020] v. 2.2.0 com o 'backend *Tensorflow* [Ten 2020] versão GPU 1.10.0.

6. Resultados

Treinamos o modelo ao longo de 3000 épocas com o *batch size* definido em 128. Utilizamos um computador com um processador Intel i7 7700, uma placa de vídeo GeForce GTX 1080ti, com 8GB de memória RAM e 1TB de capacidade de armazenamento em disco. O treinamento levou aproximadamente 50 (cinquenta) minutos. Por volta da época 2500 (dois mil e quinhentos), o modelo alcançou uma *acurácia* de 50%. Nesse ponto de equilíbrio, o discriminador não é mais capaz de reconhecer com precisão o que é uma imagem gerada pelo gerador e o que é uma imagem pertencente ao conjunto de treinamento. Na época 3000 (três mil), a perda (*loss*) do gerador foi de 16,118, a perda do discriminador foi de 8,059 e a *acurácia* foi de 50,00%. A Figura 2 mostra seis imagens geradas pelo método proposto neste estudo.



Figura 2. Seis exemplos de imagens geradas pela aplicação do modelo DCGAN neste estudo proposto.

A respeito do uso das imagens sintéticas como artifício de *data augmentation*, a Tabela 1 apresenta os parâmetros e resultados obtidos.

Tabela 1. Resultados da injeção de imagens sintéticas na base de dados do classificador.

Objetos sintéticos	Total na Classe 0	Total na Classe 1	Acurácia	Perda (Loss)
0	296	296	0,86	0,24
10	306	296	0,87	0,24
50	346	296	0,90	0,21
100	396	296	0,93	0,19
200	496	296	0,93	0,19

7. Conclusão

Além da síntese de dados, faz-se possível verificar através deste estudo que as imagens sintéticas fotorrealistas geradas com metodologia apresentada neste artigo, são artifícios importantes no tocante ao ganho de performance de modelos pré-existentes. O classificador binário - detector de *estrabismo* - teve um ganho de 7% em sua métrica de acurácia, assim como uma queda de 0,05 no valor de sua função de perda (loss).

Os autores desta pesquisa reconhecem haver oportunidades para, em trabalhos futuros, explorarem técnicas que buscam o aumento da performance em relação a capacidade da rede discriminadora do modelo generativo em reconhecer simetrias, imagens do olho esquerdo pouco compatíveis com o olho direito ainda podem ocorrer entre as imagens geradas (aspecto raramente presentes em seres humanos).

A respeito da escassez de imagens públicas de estrabismo e da oportunidade de ofertar bases de dados sintéticas: o treinamento do modelo generativo ainda necessita de uma base de dados com imagens de pessoas com *estrabismo*. Entretanto, modelos treinados com bases que possam ter licenças que impedem sua publicação podem ter seus pesos de suas redes neurais publicados sem infringir - para a maior parte das licenças existentes - nenhum termo específico e ainda assim protegendo a privacidade dos pacientes que participaram da elaboração da base de dados, uma vez que os dados sintéticos gerados pelos modelos generativos não constituem indivíduos reais. As imagens geradas são manifestações dos padrões e distribuições aprendidas pelo modelo de aprendizagem profunda.

O uso do modelo generativo, além do explicitamente demonstrado, que trata do aumento de performance de classificadores com emprego de aumento de dados, pode ser diverso. Professores da área de saúde ocular podem gerar N imagens de pessoas com estrabismo com a finalidade de usá-las como artifício didático em suas aulas, entidades de saúde podem lançar bases de dados públicas com apenas imagens sintéticas para preservarem a privacidade dos pacientes que foram voluntários na composição da base de dados 'original', entre outros usos. O modelo também pode inspirar novas aplicações em outros domínios.

Esta pesquisa contou com o apoio e dedicação incansável do Dr. Ronaldo Barcellos, que conduziu os autores deste trabalho no sentido de ajudá-los a transpor muitos obstáculos que emergiram ao longo da pesquisa - principalmente aqueles relacionados ao *estrabismo*.

Referências

- (2020). Celebrities images | kaggle. <https://www.kaggle.com/greg115/celebrities-100k>. (Accessed on 02/15/2020).
- (2020). Eye disease dataset | kaggle. <https://www.kaggle.com/kondwani/eye-disease-dataset>. (Accessed on 11/02/2019).
- (2020). Home - keras documentation. <https://keras.io/>. (Accessed on 01/27/2020).
- (2020). Keras-gan/dcgan.py at master · eriklindernoren/keras-gan. <https://github.com/eriklindernoren/Keras-GAN/blob/master/dcgan/dcgan.py>. (Accessed on 01/27/2020).
- (2020). Losses - keras documentation. <https://keras.io/losses/>. (Accessed on 01/27/2020).
- (2020). Metrics - keras documentation. <https://keras.io/metrics/>. (Accessed on 01/27/2020).
- (2020). Optimizers - keras documentation. <https://keras.io/optimizers/>. (Accessed on 01/27/2020).
- (2020). Python.org. <https://www.python.org/>. (Accessed on 01/27/2020).
- (2020). Tensorflow. <https://www.tensorflow.org/>. (Accessed on 01/27/2020).
- Abrahamsson, M., Fabian, G., and Sjöstrand, J. (1986). Photorefraction: a useful tool to detect small angle strabismus. *Acta ophthalmologica*, 64(1):101–104.
- Calimeri, F., Marzullo, A., Stamile, C., and Terracina, G. (2017). Biomedical data augmentation using generative adversarial neural networks. In *International conference on artificial neural networks*. Springer.
- Chen, Z., Fu, H., Lo, W.-L., and Chi, Z. (2018). Strabismus recognition using eye-tracking data and convolutional neural networks. *Journal of healthcare engineering*, 2018.
- De Almeida, J. D. S., Silva, A. C., De Paiva, A. C., and Teixeira, J. A. M. (2012). Computational methodology for automatic detection of strabismus in digital images through hirschberg test. *Computers in biology and medicine*, 42(1):135–146.

- De Almeida, J. D. S., Silva, A. C., Teixeira, J. A. M., Paiva, A. C., and Gattass, M. (2015). Computer-aided methodology for syndromic strabismus diagnosis. *Journal of digital imaging*, 28(4):462–473.
- Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., and Greenspan, H. (2018). Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification. *Neurocomputing*, 321:321–331.
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4):193–202.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680.
- Hwang, S.-K. and Kim, W.-Y. (2006). A novel approach to the fast computation of zernike moments. *Pattern Recognition*, 39(11):2065–2076.
- Ingram, R. (1977). Refraction as a basis for screening children for squint and amblyopia. *British Journal of Ophthalmology*, 61(1):8–15.
- Kaakinen, K. (1979). A simple method for screening of children with strabismus, anisometropia or ametropia by simultaneous photography of the corneal and the fundus reflexes. *Acta ophthalmologica*, 57(2):161–171.
- Kelleher, J. D. (2019). *Deep Learning*. MIT press.
- Kim, H.-J. and Kim, W.-Y. (2008). Eye detection in facial images using zernike moments with svm. *ETRI journal*, 30(2):335–337.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551.
- Lopes, A. C. (2006). *Diagnóstico e tratamento*, volume 1. Editora Manole Ltda.
- Lu, J., Fan, Z., Zheng, C., Feng, J., Huang, L., Li, W., and Goodman, E. D. (2018). Automated strabismus detection for telemedicine applications. *arXiv preprint arXiv:1809.02940*, page 2.
- Radford, A., Metz, L., and Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.
- Santos, J. S. and Silveira, I. F. (2020). Generating photorealistic images of people’s eyes with strabismus using deep convolutional generative adversarial networks.
- Valente, T. L. A., de Almeida, J. D. S., Silva, A. C., Teixeira, J. A. M., and Gattass, M. (2017). Automatic diagnosis of strabismus in digital videos through cover test. *Computer methods and programs in biomedicine*, 140:295–305.