

A complexidade estatística em redes de aprendizagem profunda como critério de escolha e otimização em problemas de imagens médicas

Pesquisador Principal:

Luiz Otavio Murta Junior

Resumo

Desde o seu renascimento, o aprendizado profundo (deep learning - DL) tem sido amplamente utilizado em várias tarefas de imagens médicas e alcançou um sucesso notável em muitas aplicações de imagens médicas, impulsionando-nos assim para a chamada era da inteligência artificial (IA). Sabe-se que o sucesso da IA é atribuído principalmente à disponibilidade de grandes conjuntos de dados (big data) com anotações para uma única tarefa e aos avanços na computação de alto desempenho. No entanto, a imagem médica apresenta desafios únicos que confrontam as abordagens de DL e possuem evoluções estruturais e dinâmicas em seus modelos ainda pouco conhecidos. Neste projeto de pesquisa, vislumbra-se validar um instrumento para o estudo de melhores arquiteturas de rede, bem como a análise de evolução estrutural e dinâmicas para diferentes problemas em imagens médicas. As aplicações a serem estudadas serão segmentação e apoio ao diagnóstico por imagens médicas por DL baseada em imagens. A ferramenta matemática e estatística proposta para a estimativa da complexidade estatística a ser aferida sobre o padrão de pesos da rede será a complexidade LMC. A complexidade estatística LMC é estimada como o produto da entropia de Shannon pela distância estatística do equilíbrio será testada e validada como uma métrica de auto-organização da rede enquanto comparada com a eficiência do DL. Além disso, as ferramentas de análises dinâmica e estrutural que também serão investigadas são respectivamente entropia amostral (**SampEn**), entropia amostral bidimensional (**SampEn_{2D}**), bem como suas aferições em múltiplas escalas (multiscale entropy – **MSE**) como estimativas de irregularidade estrutural. Com este estudo se pretende entender melhor o que ocorre internamente na rede durante o DL enquanto lida com imagens médicas e quais as melhores arquiteturas de redes profundas em aplicações específicas.

Palavras-chaves: Imagens médicas; Inteligência artificial; Aprendizagem profunda; irregularidade estrutural; Complexidade estatística; Suporte ao diagnóstico; Segmentação de imagens.

Departamento de Computação e Matemática
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto - USP

The statistical complexity of deep learning networks as choice and optimization criteria in medical imaging problems

Principal Investigator:

Luiz Otavio Murta Junior

Abstract

Since its renaissance, deep learning (DL) has been widely used in various medical imaging tasks and has achieved remarkable success in many medical imaging applications, thereby propelling us into the so-called artificial intelligence (AI) era. The success of AI is well known and mainly attributed to the availability of big data with annotations for a single task and the advances in high-performance computing. However, medical imaging presents unique challenges that confront DL approaches and have structural and dynamic evolutions in their models that are still little known. This research project envisaged the study and validation of a tool for the choice of better network architectures and the analysis of structural and dynamic evolution for different problems in medical imaging. The applications to be studied will be segmentation and support for medical imaging diagnosis based on DL. The proposed mathematical and statistical tool to estimate the statistical complexity of the network weights pattern will be measured by the LMC complexity. The LMC statistical complexity estimated as the product of Shannon entropy by the statistical distance from equilibrium, i.e., disequilibrium, will be tested and validated as a metric of self-organization of the network while comparing to DL efficiency. In addition, we will investigate sample entropy (**SampEn**), two-dimensional sample entropy (**SampEn_{2D}**), and their multiscale versions (**MSE**) as the dynamic and structural analysis tools of deep network units during DL training. With this study, we intend to understand better what happens inside the network during DL while dealing with medical images and the best deep network architectures in specific applications.

Keywords: Medical images; Artificial intelligence; Deep learning; Structural irregularity; Statistical complexity; Diagnostic support; Image segmentation; Image-guided therapy.

Department of Computing and Mathematics
Faculty of Philosophy, Sciences and Letters of Ribeirão Preto -
USP

Sumário

1	INTRODUÇÃO	1
1.1	Problemas no uso da IA com imagens médicas	2
1.2	Redes de Aprendizado Profundo (DL)	3
1.3	Complexidade	4
2	OBJETO DE PESQUISA	6
3	MATERIAIS E MÉTODOS	7
3.1	Medição da complexidade	7
3.2	Complexidade LMC	8
3.3	Complexidade SampEn _{2D}	11
3.4	Repositórios de Imagens Médicas	12
3.5	Métodos de avaliação	12
4	RESULTADOS PRELIMINARES LMC	14
5	RESULTADOS ESPERADOS	16
6	CRONOGRAMA	17
7	RECURSOS DISPONÍVEIS	20
	Referências	20
	REFERÊNCIAS	21

1 Introdução

O presente projeto tem por objetivo propor e avaliar a complexidade de diferentes arquiteturas de redes DL (do inglês, *deep learning*), ou redes profundas, para diferentes aplicações de imagens médicas, observando a complexidade dinâmica e estrutural das redes durante o seu aprendizado. No contexto das arquiteturas, se sabe que as redes profundas DL possuem uma variedade de tipos, passando por camadas de redes convolucionais CNN (do inglês, *convolutional neural network*), camadas de redução de dimensionalidade (*max-pooling*), camadas densas (totalmente conectada), entre outras. Nestes termos, podemos investigar a complexidade ou irregularidade dinâmica e estrutural que é adquirida pela rede ao resolver problemas de imagens médicas.

A imagem médica explora fenômenos físicos, como luz, radiação eletromagnética, radioatividade, ressonância magnética nuclear (RM) e som para gerar representações visuais ou imagens de tecidos externos ou internos do corpo humano, ou de uma parte do corpo humano, de forma não invasiva, ou por meio de um procedimento invasivo (BEUTEL et al., 2000). As modalidades de imagem mais comumente usadas na medicina clínica incluem radiografia de raios-X, tomografia computadorizada (TC), imagem de ressonância magnética (MRI), ultrassom e patologia digital. Os dados de imagem representam cerca de 90% de todos os dados de saúde, e portanto, são uma das fontes de evidência mais importantes para análises clínicas e intervenções médicas.

A DL foi considerada uma das dez tecnologias inovadoras de 2013 (HOF, 2013). Isso se seguiu ao desafio de categorização de imagens em grande escala de 2012, que introduziu a superioridade da CNN no conjunto de dados ImageNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). Naquela época, o DL emergiu como a principal ferramenta de aprendizado de máquina nos domínios de imagem geral e visão computacional, e a comunidade de imagens médicas começou um debate sobre se o DL seria aplicável no espaço de imagens médicas. As preocupações eram devidas aos desafios descritos acima, sendo o principal desafio a falta de dados rotulados suficientes, conhecido como desafio dos dados. Várias etapas podem ser apontadas como facilitadores da tecnologia de DL no espaço de imagens médicas: em 2015-2016, as técnicas foram desenvolvidas usando transferência de aprendizagem (Transfer Learning - TL), ou o que também foi chamado de "aprendizagem de recursos não médicos" (BAR et al., 2015) para aplicar o conhecimento obtido através da resolução de um problema de origem a um problema de destino diferente, mas relacionado. Uma questão chave era se uma rede pré-treinada em imagens naturais seria aplicável a imagens médicas. Vários grupos mostraram ser este o caso (SHIN et al., 2016; GULSHAN et al., 2016); usar a rede profunda treinada com base em ImageNet e fazer o ajuste fino em uma tarefa de imagem médica foi útil para acelerar a convergência de treinamento e melhorar a precisão.

Ferramentas de inteligência artificial (IA), como a tecnologia DL, podem fornecer suporte aos médicos automatizando a análise de imagens, levando ao que podemos chamar de radiologia computacional. As ferramentas automatizadas que podem ser desenvolvidas são detecção de achados patológicos, quantificação da extensão da doença, caracterização de patologias (por exemplo, em benigna versus maligna) e ferramentas de software variadas que podem ser amplamente caracterizadas como suporte à decisão. Esta tecnologia também pode estender as capacidades dos médicos para incluir a caracterização de eventos tridimensionais e variáveis no tempo, que muitas vezes não são incluídos nos relatórios radiológicos de hoje devido ao tempo limitado e às ferramentas de visualização e quantificação limitadas.

1.1 Problemas no uso da IA com imagens médicas

A imagem médica tem várias características que influenciam a adequação e a natureza das soluções de aprendizado profundo (DL). Observe que esses traços não são necessariamente exclusivos da imagem médica. Por exemplo, imagens de satélite compartilham a primeira característica descrita a seguir com imagens médicas. As imagens médicas têm múltiplas modalidades e são densas na sua resolução de pixels, além de já ter muitas modalidades, novas modalidades, como TC espectral, são inventadas rotineiramente. Mesmo para modalidades de imagem comumente usadas, a resolução de pixel ou voxel aumentou, aumentando a densidade da informação.

Por exemplo, a resolução espacial da tomografia e ressonância magnética clínica atingiu o nível submilimétrico, enquanto os tomógrafos de hoje podem adquirir 1000–2500 cortes por indivíduo, a resolução espacial do ultrassom é ainda melhor, enquanto sua resolução temporal ultrapassa o tempo real. Como resultado, uma única imagem de patologia digital em lâmina correspondente a um único núcleo de biópsia da próstata pode facilmente ocupar 10 GB de espaço com uma ampliação de 40 vezes. No geral, bilhões de estudos de imagens médicas são realizados por ano em todo o mundo, e esse número está crescendo.

Os rótulos associados às imagens médicas são esparsos e ruidosos. Rotular ou anotar uma imagem médica é demorado e caro. Além disso, diferentes tarefas requerem diferentes formas de anotação, o que cria o fenômeno da dispersão do rótulo. Devido à experiência e às condições variáveis, as inconsistências de rotulagem entre usuários e intrausuários são altas (ZOU et al., 2004), e os rótulos devem ser considerados ruidosos. O estabelecimento de padrões ouro para a rotulagem de imagens permanece uma questão em aberto.

Em 2017-2018, o aumento de dados sintéticos surgiu como uma segunda solução para processar conjuntos de dados limitados. Obviamente, o aumento clássico é um componente vital de qualquer treinamento de rede. Ainda assim, as principais questões a serem abordadas eram se era possível sintetizar dados médicos usando esquemas, como modelagem generativa, e se os dados sintetizados serviriam como exemplos médicos viáveis e iriam, na prática,

aumentar o desempenho da tarefa médica em questão. Vários trabalhos em vários domínios demonstraram que esse era, de fato, o caso. Por exemplo, em (FRID-ADAR et al., 2018), o aumento da imagem sintética com base na rede adversária geradora (GAN) mostrou gerar amostras de imagem da lesão que não foram reconhecidas como sintéticas pelos radiologistas especialistas e aumentou o desempenho da CNN na classificação de lesões hepáticas. GANs, codificadores variacionais e variações ainda estão sendo explorados e avançados em trabalhos recentes.

Uma das contribuições críticas que emergiram da comunidade de imagens médicas para segmentação de imagens foi a arquitetura U-Net (RONNEBERGER; FISCHER; BROX, 2015). Projetada inicialmente para segmentação de células microscópicas, a U-Net provou aprender recursos práticos de forma eficiente e robusta para muitas tarefas de segmentação de imagens médicas.

1.2 Redes de Aprendizado Profundo (DL)

As redes neurais profundas têm uma capacidade de modelo maior e uma capacidade de generalização mais robusta do que as redes neurais superficiais. Modelos profundos treinados em bancos de dados anotados em grande escala para uma única tarefa alcançam desempenhos excepcionais, muito além dos algoritmos tradicionais ou mesmo da capacidade humana. A partir do AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), houve uma tendência de pesquisa para aprofundar as redes, como representado por VGGNet (SIMONYAN; ZISSERMAN, 2015), Inception Net (SZEGEDY et al., 2015) e ResNet (HE et al., 2016). Conexões de salto tornam uma rede profunda mais treinável do que DenseNet (HUANG et al., 2017) e U-Net (RONNEBERGER; FISCHER; BROX, 2015). A U-net foi proposta pela primeira vez para lidar com a segmentação, enquanto as outras redes foram desenvolvidas para a classificação de imagens. A supervisão profunda (LEE et al., 2015) melhora ainda mais o poder discriminativo.

Mecanismos de atenção adversários (GAN) Goodfellow et al. (GOODFELLOW et al., 2020) propõem acompanhar um modelo gerador com um discriminador que diz se uma amostra é da distribuição do modelo ou da distribuição de dados. Redes profundas representam o gerador e o discriminador, e seu treinamento é feito por meio de uma otimização minimax. A aprendizagem adversária é amplamente usada em imagens médicas (YI; WALIA; BABYN, 2019), incluindo reconstrução de imagens médicas (WANG et al., 2019), aprimoramento da qualidade de imagem (DEWEY et al., 2019) e segmentação (YANG et al., 2017). O mecanismo de atenção (XU et al., 2015) permite a descoberta automática de “onde” e “o que” focar ao descrever o conteúdo da imagem ou tomar uma decisão holística. O aperto e a excitação (HU; SHEN; SUN, 2018) podem ser considerados um canal de mecanismo de atenção. Atenção é combinada com GAN em (ZHANG et al., 2019) e com U-Net em (OKTAY et al., 2018).

Pesquisa de arquitetura neural (NAS) e design leve: NAS (ELSKEN; METZEN; HUTTER, 2019) tem como objetivo projetar automaticamente uma arquitetura de rede profunda para alto desempenho voltada para uma determinada tarefa. Zhu et al. (ZHU et al., 2019) aplicar NAS com sucesso à segmentação volumétrica de imagens médicas. Por outro lado, o design leve (HOWARD et al., 2017; ZHANG et al., 2018), por outro lado, visa projetar a arquitetura para eficiência computacional em dispositivos com recursos limitados, como telefones celulares, mantendo a precisão. Para lidar com rótulos esparsos e barulhentos, precisamos de abordagens DL que sejam eficientes em relação a anotações. Assim, a ideia principal é aproveitar o poder e a robustez da capacidade de representação de recursos derivada de modelos e dados existentes, mesmo que os modelos ou dados não sejam necessariamente do mesmo domínio ou para a mesma tarefa, e se adaptar.

1.3 Complexidade

O conceito de complexidade estatística decorre da teoria do caos onde, para vários sistemas não lineares existe transição entre regime ordenado, regular, e não caótico para o regime caótico, onde a aleatoriedade absoluta rege o sistema. A complexidade estatística LMC conforme definida mais adiante mais cuidadosamente é composta como o produto entre a entropia de Shannon e a distância do equilíbrio, ou desequilíbrio (FELDMAN; CRUTCHFIELD, 1998). A complexidade estatística inspira o conceito de complexidade fisiológica, através do qual Goldberger et. al., (GOLDBERGER, 1996; GOLDBERGER; PENG; LIPSITZ, 2002; COSTA; GOLDBERGER; PENG, 2005) relacionam a complexidade ao bom funcionamento dos sistemas fisiológicos autorregulados, o que permite várias funções essenciais à vida. Além disso, os trabalhos de Goldberger et. al., associam os mecanismos de doença e envelhecimento de sistemas fisiológicos com a queda de complexidade. Portanto, a complexidade estaria ligada à capacidade dos seres vivos de se adaptarem às várias situações encontradas, por outro lado, a baixa complexidade estaria relacionada às condições patológicas e envelhecimento com a falta de eficiência e adaptabilidade.

A situação de auto-organização e complexidade somente existe na região chamada de borda do caos (edge of chaos), exatamente onde se observa espaços com dimensões fractais que parece permitir a emergência de comportamentos interessantes e intrigantemente ligados à possibilidade de existência da vida. **A hipótese central aqui é a de que a complexidade em IA deep learning ou deep network está relacionada diretamente à eficiência e capacidade de adaptação, ou adaptabilidade, das redes.**

A entropia amostral **SampEn** (do inglês, *sample entropy*) e a entropia multiescala **MSE** (do inglês, *multiscale entropy*) fornecem insights sobre a complexidade das flutuações ao longo de uma gama de escalas de tempo e é uma extensão das medidas de entropia de amostra padrão descritas aqui (COSTA; GOLDBERGER; PENG, 2002; COSTA; GOLDBERGER; PENG, 2005). Como qualquer medida de taxas de entropia, o objetivo é fazer uma avaliação

da complexidade de uma série temporal. Uma das principais razões para usar uma abordagem multiescala é quando a escala de tempo de relevância na série temporal não é conhecida. Por exemplo, se olharmos para a fala, seria relevante considerar as escalas de tempo das palavras em vez de sons individuais, mas se você não tinha nenhuma ideia de que o sinal de áudio representava a fala, ou talvez mesmo qualquer ideia do conceito de fala, você não saberia qual escala de tempo seria mais informativa. Portanto, seria mais informativo olhar para um intervalo de escalas de tempo. No sinais fisiológicos, o código subjacente é desconhecido e, portanto, a escala de tempo de relevância é desconhecida (SILVA et al., 2015; SILVA et al., 2016). Recentemente, tem sido proposto e aplicado em diferentes situações a entropia multiescala bidimensional MSE_{2D} para medir e analisar a irregularidade e complexidade de texturas em imagens (SILVA et al., 2018).

2 Objeto de pesquisa

Diante do exposto, o presente projeto se distingue por avançar por além das avaliações das aplicações, analisando a evolução da complexidade estrutural e dinâmica de rede durante o DL nestas diferentes aplicações. Entende-se por complexidade estrutural, no contexto de redes DL, o nível mensurável de auto-organização dos padrões de pesos sinápticos adquiridos durante o treinamento.

A hipótese que se pretende investigar é a de que as redes profundas convergem para um estado de auto-organização dos padrões de pesos (sinápticos) durante o processo de aprendizado supervisionado.

A complexidade pode ser vista como forma de medir o grau de auto-organização da rede DL, pode qualificar o treinamento, permitindo a escolha de diferentes técnicas de aprendizado DL e arquiteturas de rede adequadas a diferentes aplicações em imagens médicas. Cabe ressaltar que é previsto que a tendência de ganho de complexidade durante o aprendizado seja maior quando tratamos de imagens médicas comparativamente ao tratamento de imagens ordinárias como fotografia. Este aspecto comparativo entre imagens fotográficas ordinárias e imagens médicas em diferentes modalidades será um dos pontos do estudo dentre as diferentes investigações possíveis neste contexto.

3 Materiais e Métodos

O presente projeto pretende estudar a complexidade intrínseca de redes DL usando as linguagens `python` e `C++`, apresentando para as redes DL imagens médicas de diferentes modalidades, utilizando as ferramentas TensorFlow e pyTorch para a construção das redes e treinamento. As diferentes arquiteturas que serão testadas em diferentes conjuntos de modalidades diferentes de imagens médicas com a finalidade de mensurar a qualidade da rede DL e do treinamento através das medidas aqui colocadas, i.e., a LMC e a **SampEn**.

O TensorFlow é um sistema de aprendizado de máquina que opera em larga escala e em ambientes heterogêneos. O TF usa grafos de fluxo de dados para representar a diferentes etapas de computação, os estados compartilhados e as operações que alteram esse estado. Ele mapeia os nós de um grafo de fluxo de dados em muitas máquinas em um cluster e em uma máquina em vários dispositivos computacionais, incluindo unidades centrais de processamento (central processing units – CPUs) de múltiplos núcleos, unidades gráficas de processamento (graphic processing units – GPUs) de propósito geral (GPGPUs) e unidades de processamento de tensores (tensor processing units - TPUs) ([ABADI et al., 2016](#)). Essa arquitetura oferece flexibilidade no desenvolvimento de aplicações, avaliações e pesquisas na área de aprendizado de máquina e inteligência artificial, o TF permite o desenvolvimento experimental de novas otimizações e algoritmos de treinamento. O TF é compatível com uma variedade de aplicativos, com foco no treinamento e inferência em redes neurais profundas. Vários serviços do Google usam o TF na produção dos seus serviços, o TF foi lançado como um projeto de código aberto e se tornou amplamente usado para pesquisas de aprendizado de máquina e inteligência artificial aplicado a imagens médicas ([LEE; KIM, 2018](#); [COY et al., 2019](#)).

3.1 Medição da complexidade

Nestes termos, e tomando como exemplo a rede DL ilustrada na Figura 1, vamos inicialmente mostrar que as medidas de complexidade, i.e., complexidade LMC, **SampEn** e MSE aplicadas nas matrizes ou tensores de pesos da rede, podem ser utilizados como marcadores de qualidade das arquiteturas e treinamentos das redes DL. Então fazer a aplicações com os objetivos de classificação (diagnóstico), segmentações, ou suporte a tratamento como exemplos, a fim de testar a robustez do método.

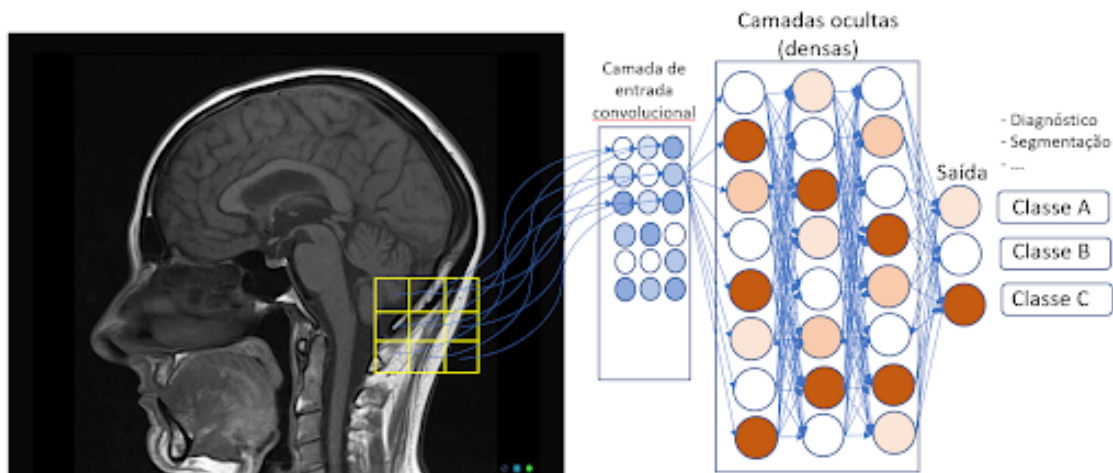


Figura 1 – Ilustração exemplificando uma possível arquitetura simples de uma rede profunda com uma camada de entrada convolucional (CNN – convolutional neural network), seguida por camadas ocultas densas ou completamente conectadas, e camada de saída com múltiplas classes. Nesta rede multicamadas estruturada espacialmente, assim como a imagem, o peso de cada nó (peso sináptico) é ajustado durante o processo de treinamento ou otimização da rede. As cores na ilustração representam as atividades em cada nó da rede em um determinado instante.

3.2 Complexidade LMC

Uma medida de complexidade estatística baseada em uma descrição probabilística de sistemas físicos e informacionais proposta por Lopez-Ruiz et. al. (LÓPEZ-RUIZ; MANCINI; CALBET, 1995), i.e., a chamada complexidade LMC, incorpora as principais características da noção intuitiva de tal medida para a qualidade de uma rede, ou da sua auto-organização, como uma grandeza que se contrapõe simultaneamente aos conceitos de ordem e aleatoriedade. Pode ser aplicado a muitas situações físicas e informacional a diferentes descrições de determinados sistemas. Além disso, o cálculo do seu valor não requer esforço computacional em muitos casos de interesse físico e computacional, sendo descrita como o produto da entropia informacional de Shannon pela distância do equilíbrio.

Nos fundamentos mais básicos, um objeto, procedimento ou sistema é dito complexo quando não corresponde a padrões considerados simples. Isso soa como um oxímoro, mas o conhecimento comum nos diz o que é simples e complexo: sistemas simplificados ou idealizações são sempre um ponto de partida para resolver problemas científicos. A noção de “complexidade” na física (GULSHAN et al., 2016; FRID-ADAR et al., 2018) começa por considerar o cristal perfeito e o gás ideal isolado como exemplos de modelos simples e, portanto, como sistemas com “complexidade” zero. Recordemos brevemente suas principais características com “ordem”, “informação” e “equilíbrio”.

Uma estrutura cristalina perfeita é completamente ordenada e os seus componentes são organizados seguindo regras rigorosas de simetria. A distribuição de probabilidade para

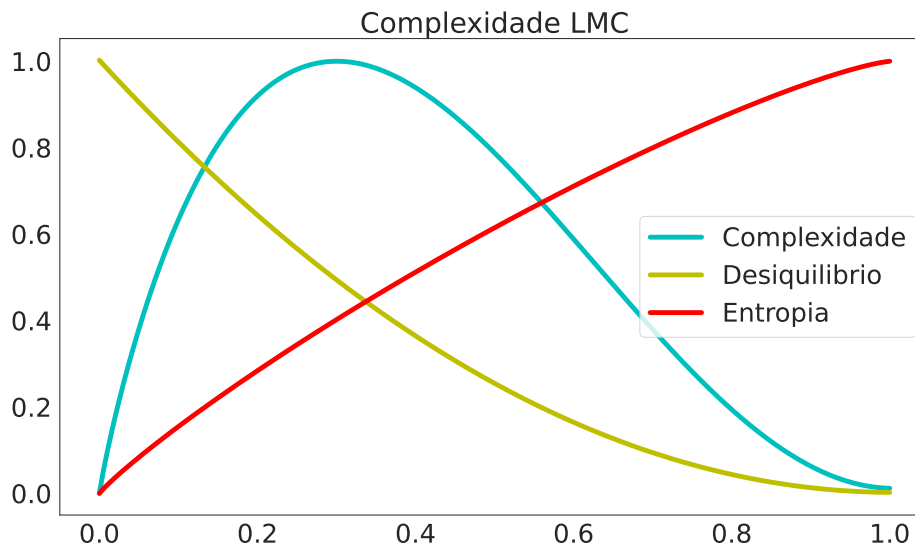


Figura 2 – Ilustração mostrando a ideia por trás do conceito de complexidade LMC. O gráfico mostra três grandezas: a entropia informacional ou informação (entropia de Shannon), a distância do equilíbrio ou desequilíbrio, e a complexidade como o produto entre informação e desequilíbrio. Tanto na situação de ordem completa como no caos completo, i.e., aleatoriedade absoluta, a complexidade é baixa, assim, a complexidade ocorre entre a ordem e o caos.

os estados possíveis desta estrutura é centrada em torno de um estado predominante de simetria perfeita. Uma pequena “informação” é suficiente para descrever a estrutura cristalina perfeita: as distâncias e as simetrias que definem sua célula elementar. As “informações” armazenadas neste sistema podem ser consideradas mínimas. Por outro lado, um sistema em caos completo isolado é completamente desordenado. O sistema pode ser encontrado em qualquer um de seus estados possíveis com a mesma probabilidade. Todos eles contribuem em igual medida para a “informação” armazenada no sistema. Tem, portanto, um máximo de “informação”. Esses dois sistemas simples são extremos na escala de “ordem” e “informação”. Segue-se que a definição de “complexidade” não deve ser feita apenas em termos de “ordem” ou “informação”.

Pode parecer razoável propor uma medida de “complexidade” adotando algum tipo de distância da distribuição equiprovável dos estados acessíveis do sistema (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). Definido dessa forma, “desequilíbrio” daria uma idéia da hierarquia probabilística do sistema. “Desequilíbrio” seria diferente de zero se houvesse estados privilegiados, ou mais prováveis, entre os acessíveis. Mas somente isso não funcionaria. Voltando aos dois exemplos com os quais começamos, vê-se facilmente que uma estrutura perfeita está longe de ser uma equidistribuição entre os estados acessíveis porque um deles é totalmente prevalecente e, portanto, o “desequilíbrio” seria máximo. Para caos completo, o “desequilíbrio” seria zero por construção. Portanto, tal distância ou “desequilíbrio” (uma medida de uma

hierarquia probabilística) não pode ser diretamente associado à “complexidade”.

A informação de Shannon ou entropia H pode continuar a ser usada como uma magnitude numa situação geral com N estados possíveis:

$$H = -k \sum_{i=1}^N p_i \log p_i \quad (3.1)$$

com k como uma constante real positiva, e p_i as probabilidades normalizadas associadas, $\sum p_i = 1$. Um sistema isolado em equilíbrio apresenta equiprobabilidade, $p_i = \frac{1}{N}$ para todo i dentro dos estados possíveis e esta é uma situação de máxima entropia.

$$H_{max} = -k \sum_{i=1}^N \frac{1}{N} \log \frac{1}{N} = -k N \frac{1}{N} \log \frac{1}{N} = k \log N \quad (3.2)$$

Se o sistema está fora do equilíbrio, a entropia H pode ser expandida em torno do seu máximo H_{max} :

$$H(p_1, p_2, \dots, p_N) = k \log N - \frac{NK}{2} \sum_{i=1}^N \left(p_i - \frac{1}{N}\right)^2 + \dots = H_{max} - \frac{NK}{2} D + \dots \quad (3.3)$$

onde surge a quantidade $D = \sum_{i=1}^N \left(p_i - \frac{1}{N}\right)^2$, que chamamos de desequilíbrio, como um tipo de distância da configuração do equilíbrio. Se a equação (3.3) for multiplicada por H , obtemos:

$$H^2 = H \cdot H_{max} - \frac{NK}{2} H D + k^2 f(N, p_i) \quad (3.4)$$

onde $f(N, p_i)$ é a entropia multiplicada pelo restante da expansão de Taylor, que apresenta a forma $\frac{1}{N} \sum_i (N p_i - 1)^m$ com $m > 2$. Se assumirmos $C = H \cdot D$,

$$C = cte \cdot H \cdot (H_{max} - H) + K \bar{f}(N, p_i) \quad (3.5)$$

com $cte^{-1} = \frac{NK}{2}$ e $\bar{f} = 2 \frac{f}{N}$. A ideia de distância do equilíbrio, i.e., desequilíbrio está claro se enxergamos que D é apenas a distância real $D \approx (H_{max} - H)$ para distâncias na vizinhança da equiprobabilidade. Em um sistema completamente caótico, temos $H \rightarrow H_{max}$ e $D \rightarrow 0$, então $C \rightarrow 0$. Ao contrário, em um sistema completamente ordenado, $H \rightarrow 0$ e $D \rightarrow 1$, mas também $C \rightarrow 0$. Esses dois sistemas são considerados exemplos clássicos de modelos simples e são extremos em uma escala de desordem (H) ou desequilíbrio (D), mas devem apresentar complexidade nula em uma medida hipotética de complexidade. Este último comportamento assintótico é verificado pela variável C na Figura 2. No lado esquerdo da figura, temos a máxima distancia do equilíbrio ou desequilíbrio e entropia mínima ou zero, indicando ordem ou previsibilidade completa, e, portanto, baixa complexidade. À medida que o sistema caminha para o equilíbrio o sistema pode passar por estágio de auto-organização, com aumento da entropia e redução do desequilíbrio, construindo estados de alta complexidade. Por outro lado, caminhando para o equilíbrio total (desequilíbrio zero) do sistema, a desordem ou entropia atinge seu patamar máximo, tornando-se totalmente imprevisível, e, portanto, complexidade zero.

3.3 Complexidade SampEn_{2D}

Outro meio de avaliar a complexidade a ser considerada neste projeto é a entropia amostral (SampEn_{2D}) estimada em diferentes escalas (MSE). A SampEn_{2D} foi definida como uma extensão da SampEn unidimensional e tem como objetivo de medir o nível de irregularidade de padrões de pesos. Brevemente, o algoritmo da SampEn_{2D} computa a possibilidade que padrões similares de tamanho m continuarão similares para tamanhos $m + 1$. Configurando a SampEn_{2D}, o tamanho de padrões (m), bem como um fator de tolerância (r) é necessário para comparar a similaridade entre dois padrões, o último refletindo como restringir pesos dentro dos padrões são considerados similares. Quanto mais alto o valor de r , mais flexível será o critério de similaridade e mais padrões similares serão encontrados pela SampEn_{2D}.

Seja $u(i, j)$ o conjunto de pesos em uma camada da rede DL estruturada em duas dimensões com largura W e altura H e $x_m(a, b)$ o subconjunto de pesos de u dentro do intervalo de (a, b) até $(a + m - 1, b + m - 1)$. Em outras palavras, $x_m(a, b)$ é uma janela quadrada de pesos de u com origem em (a, b) . As Equações de (3.6) a (3.10) definem a SampEn_{2D}

$$SampEn_{2D}(u, m, r) = -\ln \frac{U^{m+1}(r)}{U^m(r)} \quad (3.6)$$

$$U^m(r) = \frac{1}{N_m} \sum U_{i,j}^m \quad (3.7)$$

$$U_{i,j}^m = \frac{\# \text{ of } x_m(a, b) \mid d[x_m(i, j), x_m(a, b)] \leq r}{N_m - 1} \quad (3.8)$$

$$U^{m+1}(r) = \frac{1}{N_m} \sum U_{i,j}^{m+1} \quad (3.9)$$

$$U_{i,j}^{m+1} = \frac{\# \text{ of } x_{m+1}(a, b) \mid d[x_{m+1}(i, j), x_{m+1}(a, b)] \leq r}{N_{m+1} - 1} \quad (3.10)$$

onde a vai de 1 a $H - m$, b vai de 1 a $W - m$ e $(a, b) \neq (i, j)$. N_m é o número de padrões (janelas quadradas) que pode ser criada na estrutura de pesos u , para ambos os tamanhos m e $m + 1$. A função distância $d[\cdot]$ é definida na Equação (3.11).

$$d[x_m(i, j), x_m(a, b)] = \max(|u(i + k, j + k) - u(a + k, b + l)|) \quad (3.11)$$

Já a MSE_{2D} envolve a subamostragem em diferentes escalas dos sinais temporais e espaciais. A este padrão subamostrado aplica-se o algoritmo de SampEn_{2D} considerando um intervalo de escala tipicamente de 1 a 20. Com o comportamento da SampEn_{2D} em diferentes escalas se tem uma melhor ideia do tipo de processo que gerou aquele padrão, i.e., um processo estocástico, um processo que gere auto-similaridades em diferentes escalas (objetos fractais), processos oscilatórios regulares, entre outros.

3.4 Repositórios de Imagens Médicas

Com o objetivo de acessar diferentes modalidades de imagens de diferentes regiões do corpo, serão considerados repositórios públicos de imagens médicas como os detalhados abaixo. A utilização de repositórios públicos favorece também a necessária reprodutibilidade dos estudos, o que é importante em trabalhos acadêmicos.

- A coleção de imagens MedMNIST v2 (YANG et al., 2021), contendo 12 datasets de imagens 2D e 6 datasets 3D, com um total de 708.069 imagens 2D e 10.214 imagens 3D, rotuladas para um conjunto diverso de modalidades biomédicas. Este conjunto contém um Benchmark de referência e códigos para reproduzir os resultados do artigo de referência.
- O arquivo de imagens de radiologia de raios-X torácicos do instituto norte americano de saúde (National Institutes of Health – NIH) que contém 100.000 imagens, dados clínicos associados, anotações, e diagnósticos (WANG et al., 2017);
- O arquivo de imagens de cânceres (The Cancer Imaging Archive – TCIA) também do NIH (CLARK et al., 2013), com imagens do consórcio de dados pulmonares LIDC (do inglês, *Lung Image Database Consortium*), banco de dados de imagens de referência para avaliar a resposta à terapia de câncer de pulmão (RIDER) (ARMATO et al., 2008), imagens de ressonância magnética (MRI) de mama, tomografias de emissões de pósitrons (PET) / tomografias computadorizadas (CT) pulmonares, neuroimagens MRI cerebrais, colonografia CT, colonoscopia virtual (CV).
- A série de estudos de imagens de acesso aberto OASIS (do inglês, *Open Access Series of Imaging Studies*) (KLEESIEK et al., 2016) é um repositório que armazena conjuntos de dados de neuroimagem do cérebro disponíveis livremente para a comunidade científica, com 1098 indivíduos, 2168 exames de MRIs, 1608 exames de PET.

3.5 Métodos de avaliação

Para avaliar a complexidade, sua evolução e impacto em relação as redes neurais é necessário definir uma metodologia tanto sobre como será feita a medição da complexidade, quanto sobre quais redes e estruturas arquitetônicas serão avaliadas. Uma parte do cronograma será dedicada para estabelecer melhor os três critérios definidos abaixo:

Datasets Criação de um conjunto de problemas a serem resolvidos com redes neurais, e.g., classificação e segmentação usando os repositórios da seção 3.4, que servirão como referência de uma aplicação prática completa, i.e., treinamento da rede com dados e posterior validação e inferência com outro conjunto de dados para estabelecer a performance da rede.

- Redes** Definição das estruturas de rede a serem estudadas, com foco nos blocos principais, desde os modelos primitivos como blocos convolucionais e densos (*fully connected*), até módulos de arquiteturas consagradas, e.g., módulo Inception (SZEGEDY et al., 2015), módulo Residual learning (HE et al., 2016), self-attention (VASWANI et al., 2017). Como também efeitos em partes separadas das redes, e.g., comparar as partes de contração e expansão de uma U-Net (RONNEBERGER; FISCHER; BROX, 2015), comparar a complexidade na parte convolucional e densa de uma rede como AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012)
- Métricas** Definição das implementações das métricas de complexidade usadas, e.g., LMC e SampEn, junto com os meta-parâmetros de cada uma, e.g., número de bins nos histogramas ou definição da função de probabilidade que define o desequilíbrio da LMC, tamanho da janela utilizada na SampEn.

4 Resultados preliminares LMC

A Figura 3 mostra a evolução da complexidade LMC e as suas componentes, i.e., desequilíbrio e entropia. O comportamento da complexidade durante o treinamento indica que a rede se auto-organiza em termos do seu padrão de pesos para resolver o problema de classificação das imagens. Lembrando que esta dinâmica é conduzida pelos algoritmos de otimização e dirigida pela função de perda. Estes resultados sugerem que a função de complexidade pode ajudar a função de perda na otimização da rede.

Os resultados foram feitos utilizando os scripts usados pelos autores do dataset MedMNISTv2, que podem ser encontrados em <https://github.com/MedMNIST/experiments/tree/main/MedMNIST2D>, alterando apenas para exportar os logs para o TensorBoard e com uma implementação simples da entropia LMC.

Nos estudos preliminares alguns desafios a serem estudados já foram encontrados, como a definição do que é o estado de equilíbrio das redes neurais, já que as camadas podem ser inicializadas de várias formas, e.g., distribuições lineares, normais, e inicialização Xavier (Glorot; Bengio, 2010).

Outro desafio é sobre a quantidade de *bins* utilizados no cálculo do histograma que altera bastante a resposta das funções de desequilíbrio e entropia.

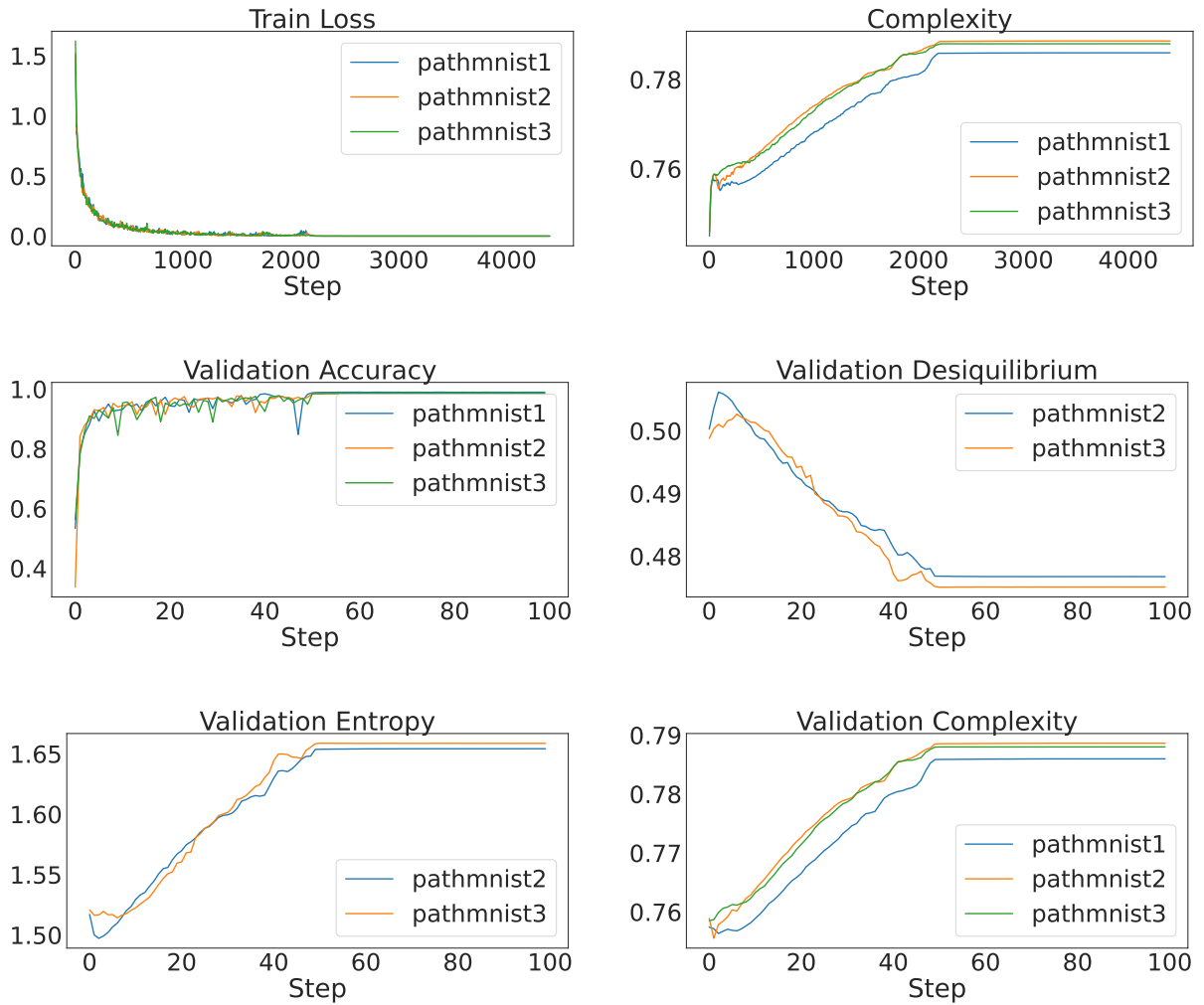


Figura 3 – Gráficos com resultados preliminares com experimento em classificação de imagens do dataset PathMNIST do conjunto MedMNISTv2 (YANG; SHI; NI, 2021). Na linha superior estão os gráficos de todas as iterações, mostrando a perda (*Train Loss*) e a complexidade (*Complexity*). As duas linhas inferiores são gráficos por época usando o conjunto de validação, mostrando a acurácia (*Accuracy*), desiquilíbrio (*Desiquilibrium*), entropia (*Entropy*) e a complexidade (*Complexity*). Os gráficos da complexidade (direita na linha superior e central) indicam que a evolução é de crescimento durante o treinamento.

5 Resultados Esperados

O presente projeto irá contribuir para entender o comportamento das chamadas redes de aprendizado profundo quando aplicada a imagens médicas quanto ao seu poder de auto-organização de padrões de pesos sinápticos. O método proposto poderá ser usado como critério para a escolha das melhores arquiteturas de redes DL e métodos de treinamento para variadas aplicações em imagens médicas. A referida avaliação em diferentes contextos e modalidades de imagens médicas levará às melhores arquiteturas de redes para as diferentes aplicações no âmbito de imagens médicas. Portanto, pode vislumbrar a produção de artigos em revistas relevante para a especialidade e patente sobre o método. Pode-se prever também a formação de doutores e mestres na especialidade de inteligência artificial em imagens médicas.

Espera-se que ao término da pesquisa seja possível treinar redes com um foco maior em desenvolver a complexidade, e não somente na redução do erro do treinamento. Isto pode levar, caso a hipótese seja válida, a redes mais eficientes e precisas, além de reduzir os recursos, e.g., o tempo, usados no treinamento das redes.

Cabe notar que o impacto desta pesquisa não fica restrito somente ao contexto de imagens médicas. Qualquer rede artificial usada em IA poderá beneficiar destes resultados, de forma maior ou menor, de acordo com sua arquitetura interna e como a complexidade afeta sua performance.

6 Cronograma

O projeto pode ser dividido em duas etapas, a saber, demonstração de validade das medidas de complexidade com índice de qualidade das redes DL; e dos seus treinamentos e aplicação e validação do índice de qualidade em diferentes contextos de arquiteturas, modalidade de imagens médicas, e finalidades de aplicações.

A primeira etapa é mais extensa e servirá como base para a segunda etapa como também para estudos futuros, fornecendo as evidências científicas sobre a relação da Complexidade com a performance e treinamento da rede. Além da criação de uma infraestrutura facilitando e automatizando os testes e processamento dos resultados, e.g., criação de tabelas e gráficos sintetizando os resultados. Isto facilitará entender o impacto de cada mudança, seja na arquitetura ou na implementação do código, na performance final da rede, que é o que importa para o público final, i.e., comunidade médica.

Na primeira etapa também será definido um *Benchmark*, com os melhores resultados possíveis com as tecnologias atuais, apenas observando a Complexidade. Na segunda etapa o objetivo será utilizar a Complexidade para melhorar as redes e superar os resultados do *Benchmark*.

Cabe ressaltar que durante todo o tempo serão enviados para publicação os resultados da primeira e segunda etapa, bem como os resultados secundários provenientes da aplicação das redes DL em diferentes contextos, as tarefas de escrita dos resultados, relatórios e manuscritos estão implícitas em cada tarefa.

A organização temporal do projeto segundo o exposto na Figura 4, utilizando os itens definidos abaixo:

Infraestrutura:

Datasets Definir os datasets a serem utilizados durante o estudo, com vários problemas diferentes, e.g., classificação e segmentação, como a aquisição dos mesmos, incluindo processos burocráticos que forem necessários para o uso de cada dataset, e.g., conformidade com as necessidades éticas de cada dataset.

Redes Definir as redes a serem utilizadas, abrangendo o maior número de blocos básicos apresentados na seção 3.5, sem se restringir a estes já que podem existir outras arquiteturas interessantes de serem estudadas, definindo também seções da rede que devam ser estudadas separadamente.

Benchmark Estabelecer resultados base entre as Redes e Datasets, tentando replicar os resultados da literatura, quando disponíveis, e estabelecer a melhor rede para cada tipo de problema, assim como suas medidas de performance como acurácia.

Métricas Implementar o código-fonte de cada métrica a ser utilizada, testando as mesmas em um subconjunto de Datasets e Redes, ajustado para ter uma variedade de dados e usar poucos recursos nesta fase. Adequando os meta-parâmetros de cada métrica como também resolvendo os detalhes da implementação numérica de cada uma no contexto de redes neurais e do framework utilizado (PyTorch ou TensorFlow).

Re-Avaliação Re-avaliar as escolhas feitas anteriormente, já com os resultados adquiridos nos primeiros trimestres, incluindo novos elementos que considerarmos interessantes estudar, e.g., novas arquiteturas e conjuntos de dados. Também pode ser uma oportunidade para alterar alguma implementação das métricas que se mostrem necessárias.

Validação da Complexidade:

Escolha arquitetura Avaliar, utilizando os dados coletados e conceitos aprendidos na etapa anterior, como a complexidade pode ajudar a escolher arquiteturas para os problemas. Arquitetura aqui não se restringe ao tipo principal de rede, e.g., InceptionNet ou ResNET, mas também, por exemplo, ao número de camadas e dimensões de cada camada.

Treinamento Avaliar como a Complexidade pode ser utilizada como uma informação adicional no treinamento das redes, em conjunto com a perda de treinamento que normalmente guia o processo de otimização das redes.

Aplicações:

Aplicações clínicas Avaliar como a complexidade pode ser utilizada para analisar como uma rede profunda aprende com dados de imagens clínicas nas tarefas de diagnóstico, segmentação, e correção. O objetivo aqui é selecionar algumas aplicações no contexto do Hospital das Clínicas de Ribeirão Preto, como o diagnóstico e prognóstico de doenças degenerativas cerebrais.

Suporte ao tratamento guiado por imagens (IGT) Testar o método em alguns cenários de suporte ao tratamento, mensurando o ganho clínico para os pacientes.

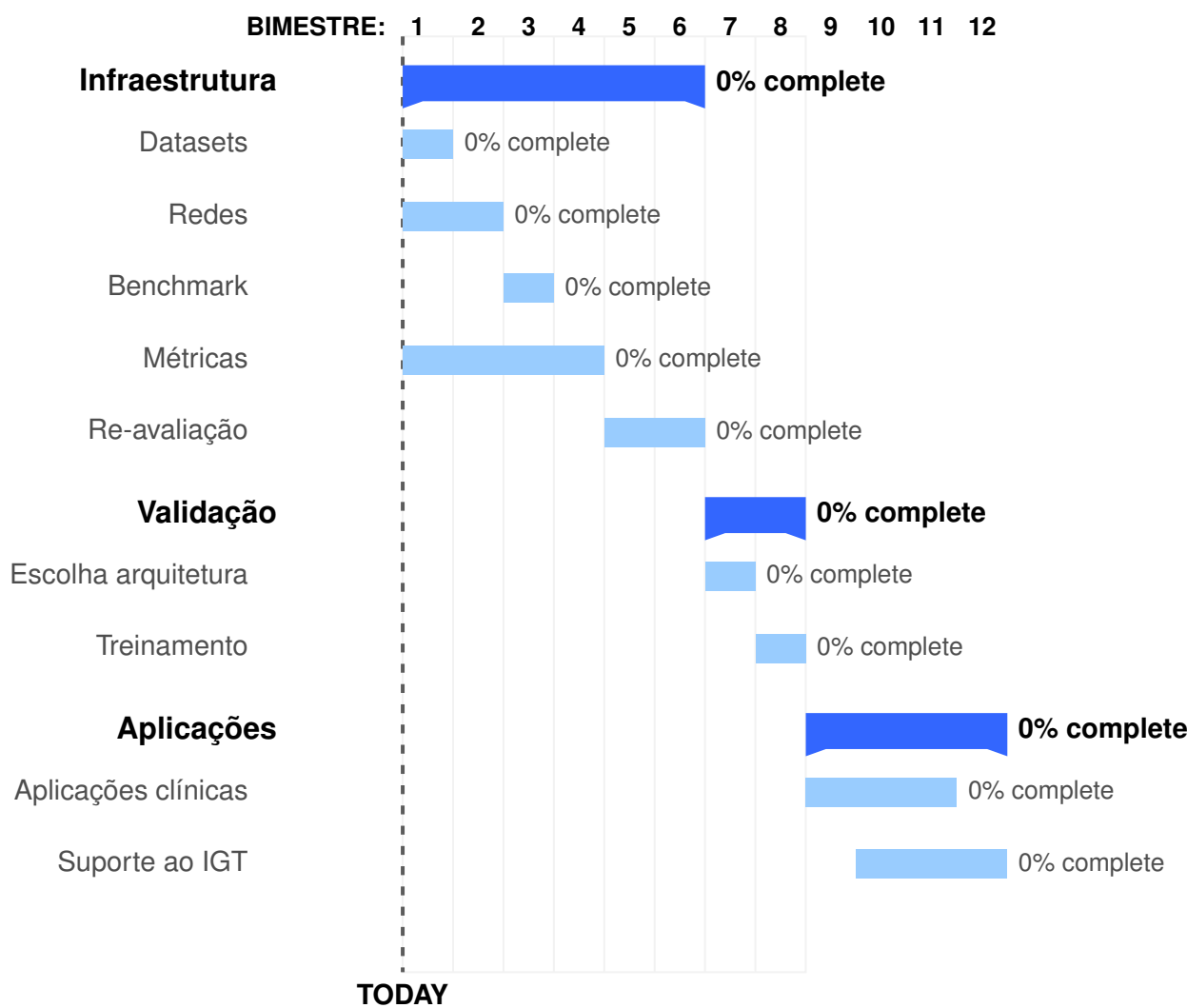


Figura 4 – Cronograma de execução do presente projeto, segundo organização trimestral.

7 Recursos Disponíveis

O grupo de pesquisa Computação em Sinais e Imagens Médicas – CSIM, coordenado pelo pesquisador proponente, dispõe de sala com computadores pessoais e um servidor de alta performance em outra sala climatizada com recursos de redundância energética (nobreaks). O servidor de alta performance computacional do grupo dispõe de uma configuração adequada e poder computacional suficiente para a execução deste projeto, com dois processadores totalizando 64 núcleos CPU de processamento, 500 GBytes de memória RAM, GPU NVidia P5000 com 2460 núcleos de processamento, e 8 TBytes de armazenamento de disco.

Referências

ABADI, M. et al. TensorFlow: a system for large-scale machine learning. In: *Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation*. USA: USENIX Association, 2016. (OSDI'16), p. 265–283. ISBN 9781931971331. Citado na página 7.

ARMATO, S. G. et al. The Reference Image Database to Evaluate Response to Therapy in Lung Cancer (RIDER) Project: A Resource for the Development of Change Analysis Software. *Clinical pharmacology and therapeutics*, v. 84, n. 4, p. 448–456, out. 2008. ISSN 0009-9236. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4938843/>>. Citado na página 12.

BAR, Y. et al. Chest pathology detection using deep learning with non-medical training. *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015. Citado na página 1.

BEUTEL, J. et al. (Ed.). *Handbook of medical imaging. Vol. 1: Physics and psychophysics*. Bellingham, Wash. (1000 20th St. Bellingham WA 98225-6705 USA): SPIE, 2000. (SPIE Press monograph, PM79). OCLC: 606655327. ISBN 9780819481184. Disponível em: <<https://stanford.idm.oclc.org/login?url=http://dx.doi.org/10.1117/3.832716>>. Citado na página 1.

CLARK, K. et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging*, v. 26, n. 6, p. 1045–1057, dez. 2013. ISSN 0897-1889. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3824915/>>. Citado na página 12.

COSTA, M.; GOLDBERGER, A. L.; PENG, C.-K. Multiscale entropy analysis of complex physiologic time series. *Physical Review Letters*, v. 89, n. 6, p. 068102, ago. 2002. ISSN 0031-9007. Citado na página 4.

COSTA, M.; GOLDBERGER, A. L.; PENG, C.-K. Multiscale entropy analysis of biological signals. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, v. 71, n. 2 Pt 1, p. 021906, fev. 2005. ISSN 1539-3755. Citado na página 4.

COY, H. et al. Deep learning and radiomics: the utility of Google TensorFlow™ Inception in classifying clear cell renal cell carcinoma and oncocytoma on multiphasic CT. *Abdominal Radiology (New York)*, v. 44, n. 6, p. 2009–2020, jun. 2019. ISSN 2366-0058. Citado na página 7.

DEWEY, B. E. et al. DeepHarmony: A deep learning approach to contrast harmonization across scanner changes. *Magnetic Resonance Imaging*, v. 64, p. 160–170, dez. 2019. ISSN 1873-5894. Citado na página 3.

ELSKEN, T.; METZEN, J. H.; HUTTER, F. Neural architecture search: a survey. *The Journal of Machine Learning Research*, v. 20, n. 1, p. 1997–2017, jan. 2019. ISSN 1532-4435. Citado na página 4.

- FELDMAN, D. P.; CRUTCHFIELD, J. P. Measures of statistical complexity: Why? *Physics Letters A*, v. 238, n. 4, p. 244–252, fev. 1998. ISSN 0375-9601. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0375960197008554>>. Citado na página 4.
- FRID-ADAR, M. et al. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, v. 321, p. 321–331, dez. 2018. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231218310749>>. Citado 2 vezes nas páginas 3 e 8.
- GLOROT, X.; BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. In: . JMLR Workshop and Conference Proceedings, 2010. p. 249–256. ISSN 1938-7228. Disponível em: <<https://proceedings.mlr.press/v9/glorot10a.html>>. Citado na página 14.
- GOLDBERGER, A. L. Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside. *The Lancet*, v. 347, n. 9011, p. 1312–1314, maio 1996. ISSN 0140-6736. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0140673696909484>>. Citado na página 4.
- GOLDBERGER, A. L.; PENG, C.-K.; LIPSITZ, L. A. What is physiologic complexity and how does it change with aging and disease? *Neurobiology of Aging*, v. 23, n. 1, p. 23–26, fev. 2002. ISSN 0197-4580. Citado na página 4.
- GOODFELLOW, I. et al. Generative adversarial networks. *Communications of the ACM*, v. 63, n. 11, p. 139–144, out. 2020. ISSN 0001-0782. Disponível em: <<https://doi.org/10.1145/3422622>>. Citado na página 3.
- GULSHAN, V. et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, v. 316, n. 22, p. 2402–2410, dez. 2016. ISSN 0098-7484. Disponível em: <<https://doi.org/10.1001/jama.2016.17216>>. Citado 2 vezes nas páginas 1 e 8.
- HE, K. et al. Deep Residual Learning for Image Recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 770–778. ISSN 1063-6919. ISSN: 1063-6919. Citado 2 vezes nas páginas 3 e 13.
- HOF, R. D. *Deep Learning*. 2013. Disponível em: <<https://www.technologyreview.com/technology/deep-learning/>>. Citado na página 1.
- HOWARD, A. G. et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv:1704.04861 [cs]*, abr. 2017. ArXiv: 1704.04861. Disponível em: <<http://arxiv.org/abs/1704.04861>>. Citado na página 4.
- HU, J.; SHEN, L.; SUN, G. Squeeze-and-Excitation Networks. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2018. p. 7132–7141. ISSN 2575-7075. ISSN: 2575-7075. Citado na página 3.
- HUANG, G. et al. Densely Connected Convolutional Networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2017. p. 2261–2269. ISSN 1063-6919. ISSN: 1063-6919. Citado na página 3.
- KLEESIEK, J. et al. Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. *NeuroImage*, v. 129, p. 460–469, abr. 2016. ISSN 1095-9572. Citado na página 12.

- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2012. v. 25. Disponível em: <https://papers.nips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>. Citado 4 vezes nas páginas 1, 3, 9 e 13.
- LEE, C.-Y. et al. Deeply-Supervised Nets. In: . PMLR, 2015. p. 562–570. ISSN 1938-7228. Disponível em: <https://proceedings.mlr.press/v38/lee15a.html>. Citado na página 3.
- LEE, J. H.; KIM, K. G. Applying Deep Learning in Medical Images: The Case of Bone Age Estimation. *Healthcare Informatics Research*, v. 24, n. 1, p. 86–92, jan. 2018. ISSN 2093-3681. Citado na página 7.
- LÓPEZ-RUIZ, R.; MANCINI, H. L.; CALBET, X. A statistical measure of complexity. *Physics Letters A*, v. 209, n. 5, p. 321–326, dez. 1995. ISSN 0375-9601. Disponível em: <https://www.sciencedirect.com/science/article/pii/0375960195008675>. Citado na página 8.
- OKTAY, O. et al. *Attention U-Net: Learning Where to Look for the Pancreas*. [S.l.], 2018. ADS Bibcode: 2018arXiv180403999O Type: article. Disponível em: <https://ui.adsabs.harvard.edu/abs/2018arXiv180403999O>. Citado na página 3.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: NAVAB, N. et al. (Ed.). *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing, 2015. p. 234–241. ISBN 9783319245744. Citado 2 vezes nas páginas 3 e 13.
- SHIN, H.-C. et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE transactions on medical imaging*, v. 35, n. 5, p. 1285–1298, maio 2016. ISSN 1558-254X. Citado na página 1.
- SILVA, L. E. V. et al. Multiscale entropy-based methods for heart rate variability complexity analysis. *Physica A: Statistical Mechanics and its Applications*, v. 422, p. 143–152, mar. 2015. ISSN 0378-4371. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378437114010462>. Citado na página 5.
- SILVA, L. E. V. et al. Two-dimensional multiscale entropy analysis: Applications to image texture evaluation. *Signal Processing*, v. 147, p. 224–232, jun. 2018. ISSN 0165-1684. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0165168418300513>. Citado na página 5.
- SILVA, L. E. V. et al. Multiscale entropy analysis of heart rate variability in heart failure, hypertensive, and sinoaortic-denervated rats: classical and refined approaches. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, v. 311, n. 1, p. R150–156, jul. 2016. ISSN 1522-1490. Citado na página 5.
- SIMONYAN, K.; ZISSERMAN, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In: BENGIO, Y.; LECUN, Y. (Ed.). *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*. [s.n.], 2015. Disponível em: <http://arxiv.org/abs/1409.1556>. Citado na página 3.

- SZEGEDY, C. et al. Going deeper with convolutions. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2015. p. 1–9. ISSN 1063-6919. Citado 2 vezes nas páginas 3 e 13.
- VASWANI, A. et al. Attention is All you Need. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2017. v. 30. Disponível em: <<https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>>. Citado na página 13.
- WANG, G. et al. *Machine Learning for Tomographic Imaging*. IOP Publishing, 2019. (2053-2563). ISBN 978-0-7503-2216-4. Disponível em: <<https://dx.doi.org/10.1088/978-0-7503-2216-4>>. Citado na página 3.
- WANG, X. et al. ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. In: . IEEE Computer Society, 2017. p. 3462–3471. ISBN 9781538604571. ISSN 1063-6919. Disponível em: <<https://www.computer.org/csdl/proceedings-article/cvpr/2017/0457d462/12OmNx8wTk9>>. Citado na página 12.
- XU, K. et al. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. In: . PMLR, 2015. p. 2048–2057. ISSN 1938-7228. Disponível em: <<https://proceedings.mlr.press/v37/xuc15.html>>. Citado na página 3.
- YANG, D. et al. Automatic liver segmentation using an adversarial image-to-image network. In: DESCOTEAUX, M. et al. (Ed.). *Medical Image Computing and Computer Assisted Intervention - MICCAI 2017*. Springer International Publishing, 2017. p. 507–515. ISBN 978-3-319-66179-7. Disponível em: <https://doi.org/10.1007/978-3-319-66179-7_58>. Citado na página 3.
- YANG, J.; SHI, R.; NI, B. MedMNIST Classification Decathlon: A Lightweight AutoML Benchmark for Medical Image Analysis. *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, p. 191–195, abr. 2021. ArXiv: 2010.14925. Disponível em: <<http://arxiv.org/abs/2010.14925>>. Citado na página 15.
- YANG, J. et al. MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical Image Classification. *arXiv:2110.14795 [cs, eess]*, out. 2021. ArXiv: 2110.14795. Disponível em: <<http://arxiv.org/abs/2110.14795>>. Citado na página 12.
- YI, X.; WALIA, E.; BABYN, P. Generative adversarial network in medical imaging: A review. *Medical Image Analysis*, v. 58, p. 101552, dez. 2019. ISSN 1361-8415. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1361841518308430>>. Citado na página 3.
- ZHANG, H. et al. Self-Attention Generative Adversarial Networks. In: . PMLR, 2019. p. 7354–7363. ISSN 2640-3498. Disponível em: <<https://proceedings.mlr.press/v97/zhang19d.html>>. Citado na página 3.
- ZHANG, X. et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2018. p. 6848–6856. ISSN 2575-7075. ISSN: 2575-7075. Citado na página 4.

ZHU, Z. et al. V-NAS: Neural Architecture Search for Volumetric Medical Image Segmentation. In: *2019 International Conference on 3D Vision (3DV)*. [S.l.: s.n.], 2019. p. 240–248. ISSN 2475-7888. ISSN: 2475-7888. Citado na página 4.

ZOU, K. H. et al. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic Radiology*, Elsevier, v. 11, n. 2, p. 178–189, fev. 2004. ISSN 1076-6332, 1878-4046. Disponível em: [https://www.academicradiology.org/article/S1076-6332\(03\)00671-8/fulltext](https://www.academicradiology.org/article/S1076-6332(03)00671-8/fulltext). Citado na página 2.