



VIABILIDADE DE DETECÇÃO DE CÂNCER DE MAMA ATRAVÉS DE REDE NEURAL CONVOLUCIONAL EM MAMOGRAFIAS

Gustavo Roncolato SOARES¹
Alisson Fernando Coelho do CARMO²

RESUMO: No universo feminino brasileiro o tipo de neoplasia maligna que possui a maior taxa de mortalidade é o câncer de mama. Quando há detecção precoce da doença o índice de sucesso é maior, resultando em melhora do prognóstico e, conseqüentemente, cura da doença. O presente estudo tem por objetivo analisar a viabilidade de um sistema capaz de detectar câncer de mama através de redes neurais convolucionais, classificando uma mamografia em cinco classes: não câncer; calcificação benigna; calcificação maligna; massa benigna; e massa maligna. Foi realizado processamento no banco de dados contendo 55.890 imagens, no qual ocorreu a conversão de dados em estrutura *tfrecords* para o formato imagem, que se faz necessária ao utilizar a rede neural. Após esta etapa, foi realizada a classificação das imagens nas cinco categorias mencionadas para viabilizar a realização de testes a fim de verificar a acurácia do algoritmo de *machine learning* em identificar e classificar o câncer de mama. Utilizando uma pequena partição de 10% imagens do banco de dados total para verificar os resultados iniciais apresentados nesse trabalho, foi possível obter 44% de acurácia global destacando a capacidade de agilizar a detecção precoce e rápida do câncer de mama utilizando inteligência artificial.

Palavras-chave: Câncer de Mama. Inteligência Artificial. Rede Neural Convolucional. Informática na Saúde.

1 INTRODUÇÃO

¹ Discente do 4º ano do curso de Sistemas de Informação do Centro Universitário Antonio Eufrásio de Toledo de Presidente Prudente. e-mail gustavoroncolato@hotmail.com Bolsista do Programa de Iniciação Científica.....

² Docente do curso de Sistemas de Informação do Centro Universitário Antonio Eufrásio de Toledo de Presidente Prudente. Doutor em Ciências Cartográficas pela Universidade Estadual Paulista Júlio de Mesquita Filho e-mail alisson.carmo@toledoprudente.edu.br Orientador do trabalho.

O câncer de mama, atualmente, representa a principal causa de morte por câncer em mulheres brasileiras sendo, dessa forma, um grande problema de saúde pública. O Brasil tem vivenciado altas taxas de incidência e mortalidade de câncer de mama anualmente. A previsão de novos casos em 2020, de acordo com o Instituto Nacional de Câncer (INCA), é de 66.280 mil. Apesar de tal cenário, as medidas necessárias à prevenção, ao diagnóstico e ao controle da doença não acompanham o aumento de casos da doença.

Os métodos de diagnósticos de câncer de mama sofreram grande avanço tecnológico, permitindo resultados mais assertivos. Apesar disso, o diagnóstico de câncer de mama ainda encontra barreiras como difícil acesso aos serviços de saúde, sobretudo em regiões do país menos favorecidas, como Norte e Nordeste.

A melhor maneira de detectar o câncer de mama é a mamografia, que possui uma taxa de detecção de 90%. No entanto, segundo a Sociedade Brasileira de Mastologia (SBM), a mamografia foi utilizada somente em 24% dos casos em 2017, explicando-se assim o aumento de mortalidade e o diagnóstico tardio.

A situação é ainda mais preocupante se o paciente é dependente do atendimento do Sistema Único de Saúde (SUS). Segundo um estudo feito por Traldi et al (2016), leva aproximadamente trinta e seis dias entre a primeira consulta médica na Estratégia Saúde da Família (ESF), até a data da mamografia. Após essa primeira etapa, é necessário aguardar um período de quarenta e seis dias para obter o resultado da mamografia e uma consulta com um médico especialista. Após a consulta com o especialista, é preciso aguardar mais trinta e nove dias para poder realizar a biópsia junto ao médico especialista e por fim aguardar mais trinta e três dias para ter o diagnóstico da biópsia em mãos para finalmente ocorrer o encaminhamento para o tratamento mais indicado.


Além disso, de acordo com cada região, ocorre variação na fila de espera para iniciar o tratamento após a última etapa referida acima. Segundo o INCA, o intervalo médio no Brasil é de aproximadamente cinquenta e nove dias (INCA, 2020).

Sendo assim, da primeira consulta ao início do tratamento, uma paciente do SUS precisa aguardar aproximadamente duzentos e doze dias, número muito superior ao recomendado pela Política Nacional de Atenção Oncológica (PNAO) que são de sessenta dias. Essa taxa 3,5 vezes maior que o recomendado, somada à

espera, faz com que a paciente desenvolva um quadro clínico mais acentuado da doença.

A grande incidência de câncer de mama no universo feminino, bem como tempo de espera considerável, sobretudo por pacientes dependentes do SUS, ressalta a urgência de métodos efetivos para diagnóstico de forma eficiente e rápida. Nesse cenário, técnicas baseadas no apoio computacional tem ganhado grande destaque. Em um estudo com técnicas de aprendizado profundo (*deep learning*), desenvolvido pelo Massachusetts Institute of Technology (MIT) foi obtido um nível de concordância de 90% entre os médicos especialistas com o grande benefício de que o processamento de dados leva segundos para emissão de resultados.

A computação vem sendo muito utilizada nos últimos tempos em todas as áreas, inclusive com grande penetração na área da saúde. Sendo assim, é necessário buscar ferramentas que possam também auxiliar no diagnóstico do câncer de mama. Dessa forma, o presente estudo busca desenvolver e aplicar métodos que permitam agilidade no processo diagnóstico da patologia em questão, bem como instituição precoce do tratamento, a fim de reduzir a taxa de mortalidade por câncer de mama. Neste trabalho foram aplicados algoritmos de *machine learning* utilizando um dos conceitos de Redes Neurais Artificiais (RNA) e Inteligência Artificial (IA), que denomina-se Redes Neurais Convolucionais (RNC). O termo convolucional é empregado pelo fato de tais ferramentas terem sido desenvolvidas com o objetivo de realizar a classificação de imagens.

 Neste trabalho foram utilizadas imagens de mamografias obtidas do bancos de dados abertos como o *Digital Database for Screening Mammography* (DDSM - Mammography), disponibilizado pela empresa Cancer Image Archives³. As RNCs possibilitam cálculos de funções matemáticas com base no reconhecimento de padrões previamente aprendidos e, com isso, realizam processamento espacial, auxiliando, desta forma, na agilização de diagnósticos de câncer de mama sem que haja necessidade de espera por uma biópsia.

O presente estudo tem por objetivo geral abordar sobre os possíveis avanços tecnológicos na área oncológica, mais precisamente no câncer de mama, através da análise da eficácia que uma RNC poderá alcançar. A principal hipótese desse trabalho é que partir do desenvolvimento de uma RNC, é possível realizar a

³ "The Cancer Imaging Archive." <https://www.cancerimagingarchive.net/>. Acessado em 7 set.. 2020.

detecção precoce e assertiva de câncer de mama, permitindo, dessa forma, que os tratamentos possam ser instituídos ainda na fase inicial do câncer, melhorando o prognóstico da paciente.

Para isso, foram desenvolvidos os seguintes objetivos específicos: coleta dos dados no banco de dados *DDSM-Mammography*; transformação e decodificação das informações coletadas em imagens, para que assim a inteligência artificial possa realizar os processamentos de convolução; análise de como uma RNC extrai as informações da imagem e como é realizado a categorização através das características estabelecidas; comparação com o estado da arte comprovando a eficiência das RNCs na utilização de imagens de mamografia para a detecção do câncer de mama; identificação da acurácia que os algoritmos de *machine learning* possa alcançar utilizando uma partição amostral do banco de dados proposto.

Em cada imagem foram analisadas alterações no aspecto da mama, tais como pele em casca de laranja, secreção serosanguinolenta em aréola, linfadenopatia cervical, supraclavicular e axilar, característica e consistência do nódulo, textura, perímetro, quantidade de concavidades no contorno do nódulo, simetria, dimensão fractal da lesão, calcificações, infiltrações, angiogênese e quadrante no qual se encontra.

2 Contextualização e Metodologia

Cada célula do organismo humano possui informações genéticas que regulam os processos de crescimento, divisões mitóticas, bem como tempo de vida celular e apoptose. Quando essas células passam por alterações de estrutura genética, ocorre um desequilíbrio destes mecanismos, resultando em proliferação celular desenfreada sem que haja morte celular programada.

Sendo assim, tais alterações resultam em neoplasia, a qual, quando adquire características como perda da relação entre núcleo e citoplasma, infiltração do tecido subjacente e elevada mitótica, recebe a denominação câncer. Quando uma neoplasia maligna tem início em tecido epitelial, como pele e mucosas, é chamada de carcinoma.

2.1 Câncer de Mama

As mutações descritas na Seção anterior podem ocorrer em células mamárias, resultando em câncer de mama. Este tipo de carcinoma se subdivide em três grupos: fibroadenoma (neoplasia benigna), carcinomas ductais e carcinomas lobulares (neoplasias malignas). Para a detecção do câncer de mama, são realizados diversos exames clínicos, sendo os mais frequentemente encontrados na prática clínica o autoexame, seguido por exames complementares como mamografia e ultrassonografia.

Por ser tratar de um procedimento que utiliza radiação, a mamografia é indicada para mulheres acima de 35 anos, para as pacientes abaixo dessa idade é realizado a ultrassonografia. Além disso, a mamografia apresenta baixa sensibilidade em mulheres jovens, visto grande densidade do tecido mamário e escassez de tecido adiposo, o que dificulta a visualização ao exame de possíveis alterações existentes.

A mamografia é um exame de imagem onde o mamógrafo, aparelho utilizado para realizar a mamografia, aplica pequenas porções de raios-x gerando assim a radiografia das mamas. Através dela são identificadas diversas alterações, tais como microcalcificações, assimetrias, nódulos, cistos, dentre outras lesões de mamas (INCA, Situação do Câncer de mama no Brasil, 2020).

A mamografia utiliza a escala BI-RANDS (Breast Imaging Reporting and Data System. 2014), medindo a proporção de tecido gorduroso e fibroglandular, classificando a densidade mamária da seguinte maneira: BIR-I, predominantemente gordurosa; BIR-II, contendo áreas dispersas com tecido fibroglandular; BIR-III, heterogeneamente denso e por fim BIR-IV, extremamente denso.

2.2 Informática na Saúde

A utilização de sistemas está presente em todos os ramos de atividades, nas mais diversas áreas, de forma que ela se adéqua a cada ramo de atividade para poder obter o melhor resultado possível. Uma das áreas que vem se beneficiando do avanço tecnológico é a medicina. São vários os profissionais da área da saúde que utilizam algum sistema de informação, para os auxílio em processos e rotinas, seja para tomada de decisões, na coleta e armazenamento de dados referente aos pacientes, medicamentos ou demais informações.

Na área da saúde também é utilizado sistemas especialistas, que são inteligências artificiais com o objetivo de auxiliar os profissionais da área de saúde a tomarem as decisões, como demonstrado por EL-DAHSHAN et al. (2014). Esses sistemas especialistas são desenvolvidos com o intuito de encontrar uma solução ao problema da mesma maneira que o profissional, desde que a solução do problema já tenha sido ensinada ao sistema especialista.

2.3 Redes Neurais Artificiais

Quando os algoritmos de RNAs foram projetados, Haykin (1999), a descreve como possuindo o mesmo princípio do cérebro humano, que pode ser considerado como o computador mais poderoso que existe, que realiza processamento de informações paralelas e não lineares, realizando conhecimento de padrões, percepções, e controle motor com uma qualidade maior que um computador convencional, ou seja, as RNAs, são desenvolvidas para emular a estrutura de funcionamento dos neurônios dos seres humanos.

A RNA, assim como a Rede Neural Humana, na qual foi baseada, é composta por neurônios conectados por conexões sinápticas, como pode ser visto na Figura 1. Cada conexão possui pesos diferentes, e a obtenção do conhecimento desejado ocorre, somente, com a atualização desses pesos.

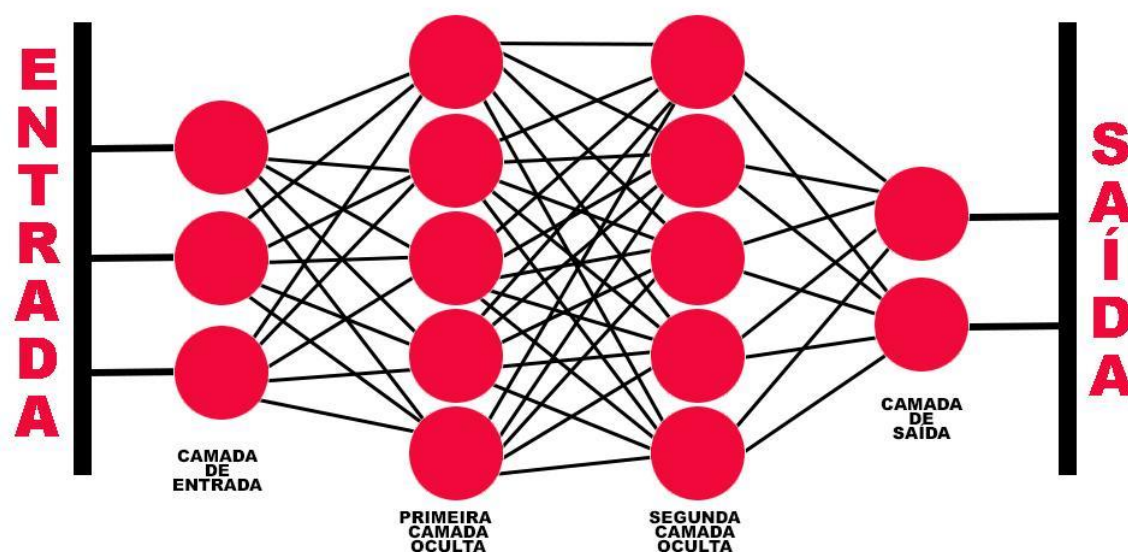


Figura 1: Rede Neural Artificial

Segundo Haykin (1999), o interesse em redes neurais artificiais tem sido motivado pelo crescente conhecimento acerca do cérebro humano. Este tem a capacidade de processar informações de forma complexa, paralela e não linear. Pode-se dizer que as RNAs procuram simular os processos de funcionamento de uma rede neural humana como: aprendizagem, generalização, associação e abstração. Basicamente o processo de funcionamento de uma rede neural artificial consiste primeiramente na etapa de aprender através de experiência e, posteriormente, utilizar o aprendizado para se fazer inferências sobre novas experiências até então desconhecidas.

As RNAs, são capazes de resolver problemas de aproximação, categorização, otimização, classificação e previsão do que irá acontecer no futuro (Rezende, 2005). Para que uma Rede Neural artificial tenha um desempenho satisfatório é necessário que ela seja tolerante a uma controlada imprecisão e possua um excelente treinamento com uma base de dados representativa, assim ela poderá mostrar uma prévia de como será o resultado final, seja para reconhecimento de imagens ou vozes, processamento de séries temporais, reconhecimento de desejo de compra, simulação de pandemia, diagnósticos médicos e entre diversas finalidades com a qual ela é utilizada.

2.4 Rede Neural Convolucional

As RNCs são métodos conhecidos há vários anos, porém que estão ganhando destaque recentemente em razão da grande capacidade computacional jamais vista. Tais algoritmos vem se demonstrando muito eficaz no processamento e análise de imagens digitais, devido ao fato de, avaliar nativamente as informações espaciais.

Desenvolvida por LeCun, et al (1998), as RNCs foram desenvolvidas de forma que suas camadas realizassem operações de forma lineares e não lineares. Sua própria nomenclatura advém do principal operador responsável por sua execução, o operador convolucional morfológico. Como pode ser visto na Figura 2, existem filtros convolucionais que extraem características de imagens. A sua utilização sobre a imagem possibilita o retorno das características desejadas, seus filtros são aplicados em várias camadas e o seu valor é assimilado pela própria rede,

determinando quais são as características relevantes para se obter a solução de saída.

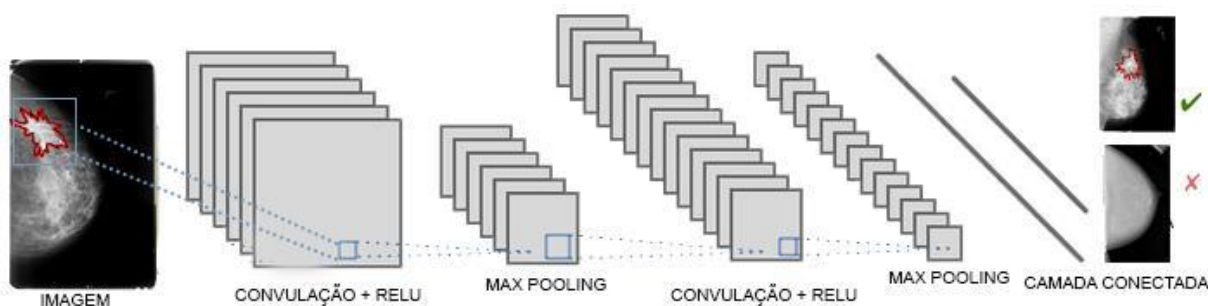


Figura 2: Exemplificação de Rede Neural Convolucional

As Redes neurais convolucionais atuam de forma bem semelhante as RNAs, porém os resultados relacionados a aplicações de imagens dela são bem superior, visto que ela possui a capacidade de avaliar informações espaciais, uma outra diferença é que as redes neurais artificiais utilizam vetores, já as RNCs utilizam, tanto na entrada, como na saída, estruturas.

As estruturas podem ter até três dimensões, na entrada, a quantidade de dimensões está interligada às pigmentações da imagem a ser utilizada, imagens monocromáticas utilizam duas dimensões, e imagens policromáticas utilizam três dimensões.

2.4.1 Convolução

A composição mais utilizada em uma RNCs, é a convolução, pois possui uma grande eficiência quando se é utilizada nas áreas de processamento de imagens. Cada camada é composta por diversos neurônios, cada neurônio desse utiliza uma espécie de máscara em uma determinada região da imagem. Este recurso tem o nome de elemento estruturante, conforme pode ser visto na Figura 3, no qual um elemento estruturante 3x3 é aplicado na imagem original para gerar o pixel na imagem de saída.

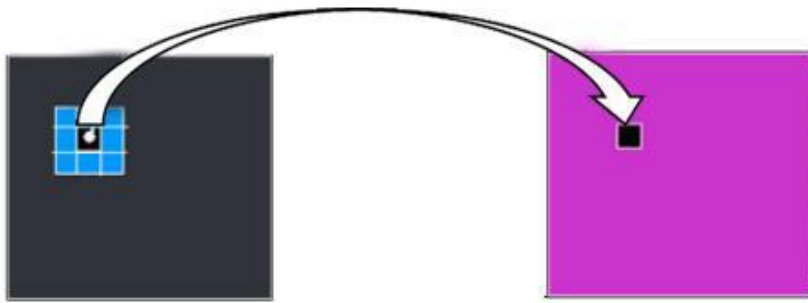


Figura 3: Processo de convolução gerando uma nova informação a partir das informações vizinhas da imagem de entrada

Na Figura 4 é possível observar as etapas de uma operação de convolução utilizando uma matriz de dimensão 5x5 pixel, com um elemento estruturante de 3x3 pixel, elemento esse que vai ser responsável por aplicar a operação de convolução. No exemplo é utilizado apenas o cálculo da primeira linha da imagem de saída.

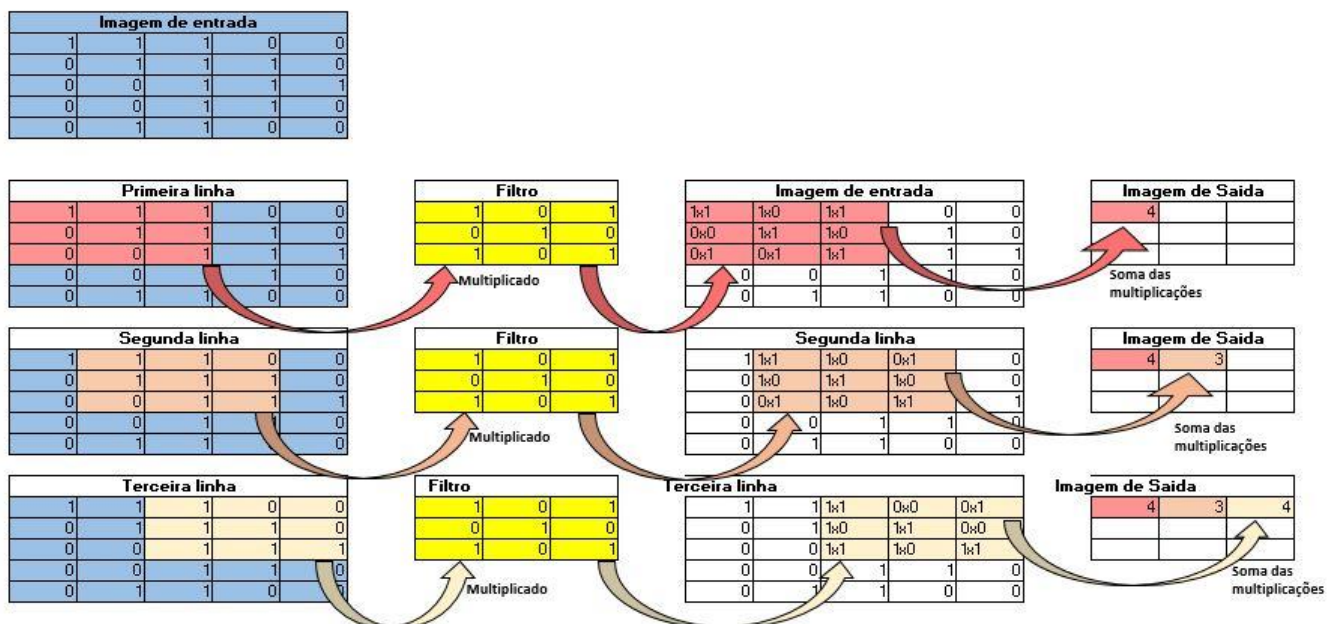


Figura 4: Exemplo de Max Pooling realizando o processo de transformação dos valores das imagens a partir do Filtro convolucional

A imagem de saída exibida na Figura 4 demonstrada acima resultou em uma matriz 3x3, podemos observar que cada valor da matriz de saída é o resultado

da soma da multiplicação que ocorre entre a imagem de entrada pelos elementos estruturantes.


A estrutura da convolução utiliza-se do mesmo princípio da RNA, onde é possível formar outras camadas com outros neurônios. Esta camada recebe uma matriz de peso, estes pesos são oriundos do treinamento que ocorre ao receber a base de dados. O fato da RNC utilizar estrutura espacial oferece uma grande vantagem para o processamento de grandes imagens, recurso esse que as redes neurais artificiais não possuem, pois, assumem uma estrutura linear de uma única dimensão.

2.7 Método

O desenvolvimento da RNC utilizada neste trabalho aconteceu no ambiente Google Colaboratory⁴, que possibilita a criação e execução de códigos escritos na linguagem Python na infraestrutura de nuvem da Google. A ferramenta não necessita de instalações e tem a sua execução em nuvem de forma online e gratuita.

Antes do início do desenvolvimento da rede neural convolucional, se faz necessária a conversão do banco de dados, visto que as imagens fornecidas pela Cancer Image Archive são disponibilizadas no formato *tfrecords* e é preciso que as imagens estejam no formato JPG ou PNG. O formato do arquivo *tfrecords* é um formato simples, utilizado para armazenar uma sequência de arquivos binários, já o formato JPG ou PNG é um formato mais utilizado nas imagens.

Em paralelo à conversão das imagens para o formato desejado (JPG) ocorreram outros dois processos: o primeiro é a divisão e classificação em cinco grupos de pastas a partir dos registros originais do *tfrecords*: mama saudável; calcificação benigna; massa benigna; calcificação maligna; e, por fim, massa maligna. O segundo processo é o armazenamento das imagens, onde será utilizado o serviço de armazenamento da própria nuvem google, o Google Drive.

 Cada pasta referente à cada grupo de classificação, após o processamento mencionado acima, armazenou cerca de 11.000 imagens; sendo assim, o banco de imagens a ser utilizado para o desenvolvimento da rede neural

⁴ "Google Colab." <https://colab.research.google.com/>. Acessado em 30 ago.. 2020.

convolucional contém mais de 55.000 imagens. Para os experimentos iniciais abordados neste trabalho, decidiu-se utilizar 10% do banco de dados, cerca de 5.500 imagens para amostragem permitindo, assim, os primeiros testes de forma ágil sem depender de longo tempo de processamento. Dessa forma, será possível verificação da análise do desempenho inicial que a rede neural alcançará com um conjunto reduzido de imagens.

A utilização de uma amostra das informações acontece pelo fato de que quando o algoritmo de *machine learning* é executado e programado com o banco de imagens completo, torna-se um processo muito custoso e demorado para o treinamento, deixando assim sua utilização quase inviável para os testes iniciais, pois ao longo do seu desenvolvimento muitos ajustes são necessários. Vale ressaltar que o longo processo é justamente o treinamento da RNC, etapa que a rede vai aprender os padrões das imagens e após esse aprendizado, qualquer classificação de uma nova imagem é realizado em tempo real, destacando o grande ganho de desempenho em comparação com a análise e diagnóstico humano.

Para a seleção das imagens integrantes do universo amostral, foram seguidas algumas diretrizes: escolha totalmente aleatória, garantia de não repetição das imagens e cada subclasse (grupo das pastas) deve ter o mesmo nível de contribuição para banco de amostra, ou seja, garantia da representação em quantidade equivalente de cada grupo.

O desempenho do modelo de uma rede neural deve sempre ser avaliado com base em dados não visualizados na etapa de treinamento, ou seja, inéditos para a RNC. Sendo assim, foi realizada a divisão da amostragem em dois conjuntos: o primeiro, denominado conjunto teste com 30% de imagens, e o segundo, conjunto treinamento com o restante do conjunto amostral. Isto ocorre pelo fato de que o conjunto de treinamento é utilizado para o aprendizado da *machine learning*. Após treinamento deste modelo, as características aprendidas são utilizadas para fazer as previsões no conjunto teste, verificando, assim, o desempenho da rede neural convolucional. Para essa finalidade, foi utilizado o método *ImageDataGenerator*, o qual permite determinar qual a porcentagem do banco amostral será utilizada para testes e qual será a outra parte a ser utilizada no treinamento.

Terminando os ajustes das imagens, iniciou-se o processo de criação da RNC, com duas camadas de conexões utilizando a progressão geométrica de base dois. Tais conexões estão interligadas por uma camada totalmente conectada,

possibilitando ocorrência dos primeiros treinamentos da *machine learning* proposta neste estudo.

Após conclusão da etapa exposta anteriormente, são iniciados os testes para verificar o resultado da rede neural convolucional. Para isso, são utilizados rótulos com verificação de classe verdade e a exibição visual, como pode ser visto na Figura 5, para facilitar a visualização da acurácia inicial obtida.

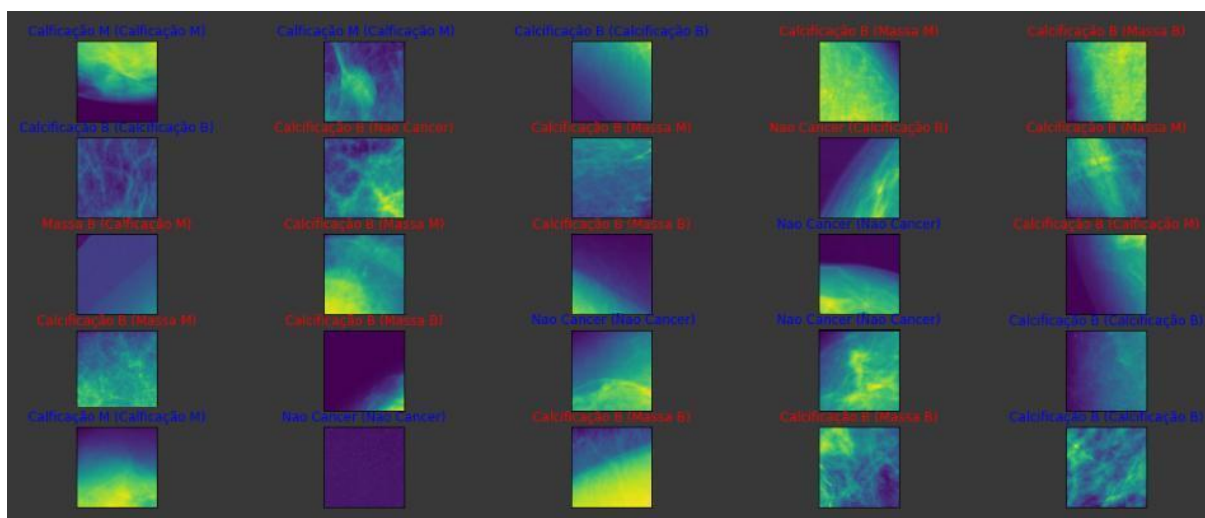


Figura 5: Demonstração das camadas de saída.

As figuras plotadas com legenda na cor azul representam os acertos da rede neural convolucional. Já as figuras com legenda vermelha representam as imagens em que ocorreu o erro da rede neural apresentada.

Com o banco amostral utilizado, foi alcançada uma acurácia de 44%, ou seja, esta é a probabilidade de acerto da *machine learning* ao se deparar com uma mamografia nunca vista. É possível analisar que ocorreu *underfitting* na *machine learning*, visto que não houve nenhum acerto quando a imagem possuía massa maligna ou calcificação maligna. Isso pode ter ocorrido por alguns fatores tais como; quantidade reduzida de imagens do banco amostral; ajustes necessários nas camadas de conexões, este erro é perceptível, pois quando a rede neural convolucional não conseguia realizar a classificação correta da imagem com base nas características previamente aprendidas, retornava como sendo calcificação benigna.

3 CONCLUSÃO

É possível inferir que *machine learning* e os algoritmos de RNCs são recursos com grande potencial a serem utilizados para auxiliar os profissionais da área da saúde em possíveis condutas terapêuticas.

Os dados do banco de imagens analisado (DDSM-Mammography) se mostraram de suma importância para a rede neural convolucional. Tal conjunto de dados apresenta uma característica de grande qualidade para a utilização em ambientes de inteligência artificial. A conversão para o formato necessário, JPEG, também obteve um resultado satisfatório e uma contribuição importante para qualquer processo que utiliza redes neurais.

A rede neural convolucional desenvolvida e apresentada nesse trabalho obteve acurácia de 44% para classificação de imagens não conhecidas pela RNCs. A acurácia apresentada se encontra abaixo do esperado mas pode ser justificada pelo reduzido conjunto de imagens utilizado correspondendo a 10% do total de imagens que o banco de dados possui.

Por fim, pode-se concluir com a expectativa de aumento da acurácia da RNC quando se for utilizado todo o banco de dados, pois com mais de 55 mil imagens existem boas perspectivas para uma acurácia melhor e mais assertiva sobre o diagnóstico do câncer de mama. Mesmo diante de um cenário experimental para realização de testes rápidos, a acurácia de 44% apresentada já representa um ótimo cenário com possibilidades de oferecer maior agilidade ao diagnóstico precoce de câncer de mama a partir de um processo totalmente automático apoiado por inteligência artificial, fato que pode ser primordial para melhoria da qualidade e velocidade de atendimento nos sistemas de saúde.

REFERÊNCIAS

ASSOCIAÇÃO MÉDICA. BRASILEIRA. **Câncer no Brasil: presente e futuro.** revista [online]. 2004, vol.50, n.1, pp. 1-1. ISSN 0104-4230. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-42302004000100001. Acesso em: 28 Março de 2020.

TRALDI, MARIA CRISTINA ET AL . **Demora no diagnóstico de câncer de mama de mulheres atendidas no Sistema Público de Saúde.** Cad. saúde colet., Rio de Janeiro , v. 24, n. 2, p. 185-191, June 2016 . Disponível em: <https://doi.org/10.1590/1414-462X201600020026>. Acesso em: 28 Março de 2020

INCA, “Instituto nacional de câncer José Alencar Gomes da Silva,” **A situação do câncer de mama no Brasil: 2020**, Disponível em: https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/a_situacao_ca_mama_brasil_2019.pdf. Acesso em Março de 2020.

SBM, **Sociedade Brasileira de Mastologia**, Disponível em: <https://www.sbmastologia.com.br/noticias/mamografia-no-brasil-o-pior-cenario-dos-ultimos-cinco-anos>. Acesso em Março de 2020.

REZENDE, S. O. (Org.). **Sistemas Inteligentes: fundamentos e aplicações.** São Paulo: Monole, 2005.

Constance D. Lehman, Adam Yala, Tal Schuster, Brian Dontchos, Manisha Bahl, Kyle Swanson, Regina Barzilay. **Mammographic Breast Density Assessment Using Deep Learning: Clinical Implementation. *Radiology*, 2018; 180694 DOI:** Disponível em: <https://pubs.rsna.org/doi/full/10.1148/radiol.2018180694>. Acesso em Março de 2020.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. **Gradient-based learning applied to document recognition. *Proceedings of the IEEE***, v. 86, n. 11, p. 2278–2324, 1998.

Banco de dados do Digital Database for Screening Mammography (DDSM). [Online]. Disponível em: <https://www.kaggle.com/skooch/ddsm-mammography> Acesso em Abril de 2020.

Liu L, Deng M. **An evolutionary artificial neural network approach for breast cancer diagnosis, *Knowledge Discovery and Data Mining Third International Conference*, 2010;10(1):593-596.**

Ministério da saúde. [Online]. Disponível em: <https://saude.gov.br/gestao-dos-sus/programacao-regulacao-controle-e-financiamento-da-mac/sistemas-de-informacao-em-saude>. Acesso em Março de 2020.

HAYKIN, S., **Redes Neurais: Princípios e Prática. 2 ed.** Editora Prentice Hall, Inc, 1999.

LUGER, G. F. **Inteligência artificial: estruturas e estratégias para a resolução de problemas complexos.** (4. Ed.). Porto Alegre: Bookmann, 2004. 774 p.

Khan, S., Rahmani, H., Shah, S. A. A., and Bennamoun, M. (2018). **A Guide to Convolutional Neural Networks for Computer Vision** Morgan & Claypool, Australia, 1 edition.

C. D'Orsi and Acr, 2013 ACR BI-RADS Atlas: **Breast Imaging Reporting and Data System**. American College of Radiology, 2014.

INCA, "Instituto nacional de câncer José Alencar Gomes da Silva," **O que é o câncer, 2020**, ´ acesso em: 18/03/2020. [Online]. Disponível em: <https://www.inca.gov.br/o-que-e-cancer>. Acesso em Março de 2020.

Stopeck, Alison. **Breast Cancer**. Medscape. [Online] Disponível em: <http://emedicine.medscape.com/article/1947145-overview#aw2aab6b2b2>. Acesso em Março de 2020.

El-Dahshan, El-Sayed & Mohsen, Heba & Revett, Kenneth & M.Salem, Abdel-Badeeh. (2014). **Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm**. *Expert Systems with Applications*. 41. 5526–5545. 10.1016/j.eswa.2014.01.021.

MORAVEC, Hans. **When will computer hardware match the human brain**. *Journal of evolution and technology*, v. 1, n. 1, p. 10, 1998.