



Fusão de Dados de Raio-x com Dados Clínicos para Detectar Covid-19

X-Ray Images and Clinical Information Data Fusion for Detecting COVID-19

Fusionar datos de rayos X con datos clínicos para detectar Covid-19

Gabriel Alves Saraiva¹, Alberto de Almeida Campos Gonçalves¹, Rafael Albuquerque Pinto², Eduardo Souto³, Eulanda Miranda dos Santos³

RESUMO

Descritores: Covid-19; Aprendizagem de Máquina

Objetivo: Este trabalho investiga o uso de fusão de dados clínicos de pacientes com informações extraídas de imagens de raio-x no contexto de classificação de Covid-19. **Método:** Duas abordagens de fusão são investigadas: fusão em nível de decisão e fusão em nível de características. Os métodos de classificação utilizados são Redes Neurais Convolucionais (CNN) e algoritmos clássicos de Aprendizagem de Máquina. **Resultados:** Os experimentos mostram que a fusão de dados melhora o desempenho do sistema em relação ao uso de classificadores individuais. Além disso, mostra-se uma superioridade da fusão em nível de decisão sobre a abordagem de fusão em nível de características. **Conclusão:** A fusão em nível de decisão balanceou melhor a contribuição de cada grupo de atributos. Esse aspecto precisa ser melhor trabalhado na abordagem de fusão em nível de características.

ABSTRACT

Keywords: COVID-19; Machine Learning

Objective: This work investigates fusing clinical data with information extracted from X-ray images in the context of COVID-19 classification. **Method:** Two fusion approaches are investigated: early and late fusion. The classification methods used are the Convolutional Neural Networks (CNN) and classic Machine Learning algorithms. **Results:** The experiments show that data fusion improves the performance of the system when compared to individual classifiers. In addition, late fusion outperformed the early fusion approach. **Conclusion:** The late fusion approach established a better tradeoff in terms of the impact of features from each data source. This aspect needs to be better addressed in early fusion.

RESUMEN

Descriptores: Covid-19; Aprendizaje Automático

Objetivo: En este trabajo investigamos la utilización de técnicas de fusión de datos clínicos de pacientes con informaciones extraídas de imágenes de rayos X en el contexto de la clasificación Covid-19. **Método:** Se investigan dos enfoques: fusión de decisiones y fusión de vectores de características. Los métodos de clasificación utilizados son redes neuronales convolucionales (CNN) y algoritmos clásicos de aprendizaje automático. **Resultados:** Los experimentos muestran que la fusión de datos mejora el desempeño del sistema en comparación con las clasificaciones sin fusión. Además, la fusión de decisiones superó a la fusión de características. **Conclusión:** La fusión a nivel de decisión equilibró mejor la contribución de cada grupo de características. Este aspecto debe tratarse mejor en el enfoque de fusión a nivel de características.

¹ Aluno de Graduação do Instituto de Computação, Universidade Federal do Amazonas (UFAM), Manaus (AM), Brasil.

² Aluno de Mestrado em Informática, Universidade Federal do Amazonas (UFAM), Manaus (AM), Brasil.

³ Professor Associado do Instituto de Computação, Universidade Federal do Amazonas (UFAM), Manaus (AM), Brasil.

INTRODUÇÃO

O novo coronavírus humano, denominado Covid-19, surgiu no final de 2019. Devido à sua rápida disseminação, a Organização Mundial de Saúde (OMS) declarou estado de pandemia, e portanto, uma emergência de saúde global, no final de janeiro de 2020. Segundo Wu et al.⁽¹⁾, a Covid-19 é uma doença respiratória aguda, podendo levar à pneumonia severa e à morte. Geralmente, os relatos iniciais da infecção surgem como um quadro de pneumonia com origem desconhecida, e seus sintomas podem variar desde a ausência (assintomáticos), até aos mais comuns: febre, fadiga, tosse seca, dispnéia, coriza ou outros sintomas do trato respiratório superior. O diagnóstico da Covid-19 pode ser realizado por exame físico, exame de imagem de tomografia computadorizada e exame radiológico de tórax.

Embora o vírus seja recente, é possível identificar um número significativo de trabalhos que propõem métodos que empregam técnicas de Aprendizagem de Máquina (AM) para diagnóstico automático da Covid-19. Em sua maioria, esses trabalhos utilizam imagens como fonte de informação. Nesse caso, a Covid-19 pode ser detectada principalmente utilizando dois exames⁽²⁾: radiografia e tomografia computadorizada. É possível observar com mais detalhes os tecidos do pulmão e ter um diagnóstico mais preciso em relação à doença na tomografia computadorizada. Porém, devido ao seu alto custo, não é um processo acessível e demanda mais tempo. Em contrapartida, a radiografia é um processo rápido, com menor custo financeiro, e consequentemente, mais acessível para um número maior de pessoas⁽²⁾.

Além de imagens, informações obtidas de notas clínicas dos pacientes também têm sido utilizadas para gerar modelos de AM para detectar Covid-19⁽³⁾. Por outro lado, na prática, é comum que os médicos durante os procedimentos clínicos utilizem informações visuais (exames de imagens) em conjunto com dados clínicos para melhorar a precisão do diagnóstico. O uso de informações clínicas de pacientes tem ajudado a melhorar significativamente a tarefa de predição em outros problemas de apoio a diagnóstico médico, como o problema de detecção de câncer de pele⁽⁴⁾, por exemplo. Não é surpresa, portanto, que esse tipo de abordagem também tenha sido utilizada para o diagnóstico automático de Covid-19⁽⁵⁻⁶⁾.

Ahsan et al.⁽⁵⁾, por exemplo, desenvolveram um modelo de diagnóstico de Covid-19 que usa Perceptron Multicamadas (MLP do inglês *Multilayer Perceptron*) e Redes de Convolução (CNN do inglês *Convolutional Neural Network*) para combinar dados numéricos/categóricos com dados obtidos de imagens de raios-X de tórax. Nessa abordagem, o MLP foi usado para lidar com os dados numéricos/categóricos, enquanto a CNN foi usada para extrair características das imagens. Dados clínicos como idade, sexo e temperatura foram utilizados como atributos. Os autores mostraram que, em uma base de dados balanceada, o modelo treinado com o algoritmo Adam conseguiu distinguir entre pacientes com Covid-19 e sem Covid-19 com 96,3% de acurácia. Enquanto em uma base de dados desbalanceada, o modelo treinado com o algoritmo Rmsprop atingiu acurácia de 95,38%.

Mei et al.⁽⁶⁾ apresentam um modelo para identificação automática de Covid-19 que também combina imagens com dados clínicos, porém, esses autores empregam imagens de tomografia computadorizada de tórax. As informações clínicas utilizadas foram: histórico de viagens e exposição, contagens de leucócitos, sintomatologia, idade e sexo do paciente. Os autores empregaram uma CNN para aprender as características de tomografias de pacientes com Covid-19, em seguida, usaram MLP para classificar os pacientes com Covid-19 de acordo com as informações clínicas. O modelo misto proposto alcançou uma acurácia de 83,5%.

Tanto Ahsan et al.⁽⁵⁾ quanto Mei et al.⁽⁶⁾ exploraram a utilização de fusão de dados em nível de decisão para detecção automática de Covid-19. Essa estratégia é dividida em duas etapas: 1) treinamento de diferentes preditores e 2) fusão das decisões dos diferentes preditores em uma única decisão. A expectativa é aumentar a capacidade de predição do sistema ao combinar diferentes decisões dos membros. Nos dois trabalhos citados, os conjuntos foram compostos por apenas dois membros: um preditor treinado com os dados clínicos e outro treinado com os dados das imagens.

Entretanto, fusão de dados pode também ser realizada em nível de características. Nesse caso, atributos de diferentes fontes de informação são extraídos, combinados e então, submetidos a um método de classificação para que apenas uma predição seja fornecida. Pereira et al.⁽²⁾ investigam tanto fusão de dados em nível de característica quanto em nível de decisão no problema de detecção de Covid-19. Porém, esses autores não utilizam dados clínicos. Ao invés disso, os autores realizam uma exploração de dados apenas em imagens de raio-X de tórax considerando diferentes métodos de extração de características. As características geradas pelos diferentes descritores são então utilizadas nas abordagens de fusão em nível de características e em nível de decisão para classificar diferentes tipos de pneumonia e, consequentemente, identificar Covid-19 entre os tipos de pneumonias.

A estratégia de fusão de dados em nível de característica ainda não tem sido muito explorada na detecção automática de Covid-19. Porém, essa estratégia tem sido utilizada na detecção de outras doenças, tais como no diagnóstico de câncer de pele^(4,7) e na detecção de doenças mentais⁽⁸⁾, por exemplo. Além disso, com base em nosso levantamento de literatura, ainda não há trabalhos que empreguem fusão em nível de características para combinar dados de imagens de raio-x com dados clínicos.

Considerando o contexto apresentado, o objetivo deste trabalho é aplicar fusão em nível de características para combinar dados clínicos de pacientes com informações extraídas de imagens de raio-x para detectar casos de Covid-19 e diferenciá-los de casos de pneumonia causadas por outras doenças. Além disso, a abordagem de fusão em nível de características é comparada à abordagem de fusão em nível de decisão.

MÉTODO

Conforme descrito anteriormente, duas estratégias de

uso de CNN são investigadas neste trabalho: 1) fusão de dados em nível de característica; 2) fusão de dados em nível de decisão. Esses dois métodos são descritos nesta seção.

Fusão em Nível de Características

A utilização de dados multimodais tem se mostrado de grande ajuda para melhorar o desempenho de modelos de classificação⁽⁸⁾. Nesse método, procuramos trabalhar a fusão de características extraídas de imagens com dados clínicos, buscando uma abordagem similar a de um médico ao diagnosticar um paciente. Para fazermos a fusão desses dois tipos de dados, primeiramente extraímos as características das imagens de raio-x utilizando uma CNN. O modelo utilizado foi a VGG-16, pré-treinada com o conjunto de dados ImageNet. Seguindo o processo de transferência de aprendizado, a rede de classificação da VGG-16 foi retirada e substituída por uma rede que foi treinada para detectar Covid-19 usando dados de imagem e dados clínicos.

A VGG-16 foi escolhida para essa tarefa devido a dois motivos. O primeiro motivo é o fato de a VGG-16 apresentar uma distribuição de pesos capaz de extrair diversas características, pois foi treinada em um conjunto de dados muito grande e com diversos tipos de imagens. A segunda razão está relacionada com a quantidade de camadas relativamente pequena dessa arquitetura. Como o conjunto de dados utilizado neste trabalho apresenta poucas imagens, um número muito grande de camadas de convolução poderia ocasionar super-ajuste mais severo do modelo aos dados.

Após fazermos a extração das características das imagens, ocorre o processo de fusão. Devido à pequena quantidade de dados disponíveis para os experimentos, optamos por utilizar um método simples de fusão: a concatenação. Para tanto, as características das imagens são inicialmente organizadas em um único vetor. No nosso caso, as características geradas pelas camadas de convolução da VGG-16 produziram um vetor de tamanho 25.088. Em seguida, esse vetor é concatenado aos atributos de dados clínicos. Neste trabalho, conforme descrito na seção que fala sobre a base de dados, sete atributos clínicos são utilizados. Portanto, essa operação resulta em um vetor final de atributos de dimensão igual a 25.095. Esse vetor é utilizado na rede neural densa treinada para a classificação das instâncias.

A rede densa utilizada foi inicializada de acordo com o método inicializador Uniforme Glorot. Trata-se uma rede com apenas duas camadas totalmente conectadas. A configuração da rede é a seguinte. Primeira camada: 100 neurônios, dimensão de entrada: 25.096, função de ativação: ReLU; segunda camada: 1 neurônio, função de ativação: Sigmoidal, regularizadores de kernel: L1 e L2. A classificação feita é binária, isto é, pneumonia causada por Covid-19 vs não-Covid-19.

Fusão em Nível de Decisão

Neste tipo de fusão, o processo está dividido em três etapas: extração de características tanto das imagens de raio-x quanto das notas clínicas; treinamento de um modelo

de classificação para cada grupo de características; e uso de uma função de fusão para obter uma única decisão a partir das decisões fornecidas pelos dois modelos de classificação obtidos. Para facilitar a comparação dos resultados das duas estratégias de fusão, as informações das imagens de raio-x foram aprendidas pelo modelo VGG-16, enquanto as notas clínicas foram aprendidas por algoritmos clássicos de AM. A fusão em nível de decisão foi realizada utilizando a regra da soma.

Além de usar a VGG-16 para extração de características das imagens, a rede neural densa utilizada para classificar as imagens de raio-x manteve configurações similares ao modelo usado na estratégia de fusão em nível de características. A única diferença está na dimensão da camada de entrada, que tem 25.088, dado que aqui a fusão é feita somente em nível de decisão. Portanto, o treino da rede neural ocorreu utilizando 100 épocas e *batch size* igual a 10. A função de perda utilizada foi a *Binary Cross Entropy*, o otimizador escolhido foi o *Rmsprop*, assim como ocorreu na estratégia de fusão em nível de características.

Para os treinamentos com as notas clínicas, foram testados 7 algoritmos de classificação: *Naive Bayes* (NB), Floresta Aleatória (RF), Árvore de Decisão (DT), k-Vizinhos mais Próximos (KNN), Máquinas de Vetores de Suporte (SVM), MLP e AdaBoost. Não houve ajuste de parâmetros para esses algoritmos. Portanto, todos os métodos utilizam parâmetros padrão das bibliotecas Sklearn e Tensor Flow (utilizadas neste trabalho) durante os treinamentos com as notas clínicas.

Os resultados obtidos pelas 7 técnicas de AM clássica foram comparados e os algoritmos com melhor desempenho na tarefa de detecção de Covid-19 utilizando dados clínicos foram selecionados para serem combinados com o modelo de CNN treinado com as imagens de raio-x. A fusão entre os dois modelos é realizada por meio da função da soma, da seguinte maneira. Dada a saída de um classificador c_i fornecida como probabilidades de classe $P(w_k/y_i(x))$, a qual denota que o rótulo da instância x é w_k se o classificador c_i atribuir o rótulo y_i . A decisão da fusão é obtida por:

$$y(x) = \max_{k=1}^n \sum_{i=1}^c P(w_k | y_i(x)) \quad (1)$$

Nessa equação, c indica o número de modelos e n a quantidade de instâncias. Portanto, a classe que obtiver a maior probabilidade após a soma das probabilidades será a classe predita pelo modelo de fusão.

Base de Dados

Neste estudo foram utilizadas as imagens de raio-x e notas clínicas obtidas no repositório *COVID-19 Image Data Collection*⁽⁹⁾. A base apresenta um significativo desbalanceamento no conjunto de dados, sendo a classe Covid-19 majoritária. Embora a base possua instâncias com dados clínicos já extraídos, há muitas instâncias com valores de atributos ausentes, não sendo viável a utilização desses atributos. Para solucionar esse problema, os atributos clínicos utilizados neste trabalho foram extraídos

diretamente da seção de notas clínicas. É possível encontrar nessa seção anotações sobre o paciente feitas durante seu atendimento e estadia no hospital, nas quais podemos encontrar sintomas que os pacientes apresentavam como tosse, febre, dor de garganta, entre outros.

Após uma extensa verificação do vocabulário utilizado e da frequência com que alguns sintomas eram apresentados nas notas clínicas, foi feita a extração dos seguintes 5 atributos: febre, tosse, falta de ar, consolidação nos pulmões e opacidade nos pulmões, sendo esses dois últimos, características que foram identificadas pelos médicos nas radiografias dos pacientes. Dois atributos clínicos já extraídos que são mais amplamente representados na base de dados também foram utilizados: sexo e idade. Portanto, o vetor de atributos clínicos final é composto por 7 atributos. Os atributos nominais foram convertidos para atributos numéricos da seguinte forma: os valores 1, 0 e -1 foram atribuídos aos atributos respectivamente quando, a nota clínica indicava a presença do sintoma, ausência do sintoma ou não mencionava o sintoma.

Após a extração dos atributos, e para evitar a necessidade de fazer imputação de valores de atributos ausentes, apenas 640 instâncias foram selecionadas para compor a base de dados dos nossos experimentos. A distribuição de instâncias entre as duas classes é: 388 instâncias de Covid-19 e 252 de não-Covid-19.

Para ambos os métodos testados neste trabalho, foi utilizada validação cruzada estratificada com 5 partições para a avaliação dos modelos. Entretanto, o conjunto de dados apresenta diferentes instâncias pertencentes a um mesmo paciente, pois uma instância foi gerada a cada novo dia em que o paciente foi examinado enquanto esteve internado no hospital. Como instâncias de um mesmo paciente podem apresentar valores similares ou idênticos, nós optamos por evitar distribuir instâncias de um mesmo paciente entre as diferentes partições de dados, ou seja, um paciente representado no conjunto de treino não foi representado no conjunto de teste. Portanto, a escolha dos pacientes para cada partição de dados foi feita de forma aleatória, e todos os dados de um mesmo paciente foram utilizados na mesma partição de dados. Dessa forma, nós esperamos evitar o viés de utilizar na base de teste dados com distribuição similar ou idêntica aos dados do treino.

Métricas de Avaliação

Para avaliar o desempenho da metodologia proposta para detecção de Covid-19, foram utilizadas as métricas de acurácia (AC), precisão (PR), sensibilidade (SE), especificidade (ES) e F1-score ($F1$) representadas nas equações exibidas na tabela 1 a seguir. Os valores TP e FP referem-se a casos verdadeiros-positivos (o classificador detecta um caso Covid-19 corretamente) e casos falsos-positivos (o classificador detecta um caso não-Covid-19 como Covid-19), respectivamente. Os valores TN e FN referem-se a casos verdadeiros-negativos (o classificador detecta um caso sem Covid-19 corretamente) e casos falsos-negativos (o classificador detecta um caso de Covid-19 como não-Covid-19), respectivamente.

RESULTADOS E DISCUSSÃO

Os experimentos foram divididos em três séries: 1) fusão em nível de decisão; 2) fusão em nível de características; e 3) comparação entre as duas abordagens.

Fusão em Nível de Decisão

Conforme mencionado anteriormente, 7 técnicas de AM foram treinadas somente com os dados clínicos e os resultados comparados. A tabela 2 mostra essa comparação entre esses classificadores clássicos de AM sem fusão. O melhor resultado de cada métrica está em negrito.

Essa tabela mostra que os classificadores Naive Bayes, RF, MLP e SVM obtiveram os melhores resultados com os dados clínicos. Dessa forma, esses classificadores foram selecionados para realizar a fusão com a CNN que foi treinada com as imagens de raio-x. A tabela 3 mostra os resultados obtidos com a fusão dos classificadores, além de mostrar os resultados obtidos pela CNN utilizando somente imagens de raio-x.

É possível observar que a fusão melhorou os resultados em relação aos modelos treinados individualmente. O melhor modelo foi a fusão VGG16+MLP, que obteve acurácia de 75,27% e F1-score de 82,62%. Entretanto, os resultados também mostram, por meio dos valores de especificidade obtidos, que os modelos tiveram uma grande dificuldade em classificar a classe negativa. Por exemplo, a fusão (VGG16+MLP) obteve especificidade de 44,75%. Isso pode

Tabela 1 - Métricas de avaliação

$$AC = \frac{TP + TN}{TN + TP + FP + FN} \quad PR = \frac{TP}{TP + FP} \quad SE = \frac{TP}{TP + FN}$$

$$ES = \frac{TN}{TN + FP} \quad F1 = 2 * \frac{SE * PR}{SE + PR}$$

Tabela 2 - Resultado médio das partições baseando-se nas informações clínicas.

Modelos	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Naive Bayes	68,56%	80,59%	75,32%	54,47%	77,87%
RF	62,70%	78,40%	69,61%	47,87%	73,75%
MLP	67,68%	81,69%	72,54%	56,68%	76,84%
SVM	69,46%	78,61%	82,07%	45,88%	80,30%
DT	57,21%	75,66%	59,77%	49,57%	66,78%
AdaBoost	57,16%	71,77%	59,48%	49,53%	65,07%
K-NN	56,64%	74,42%	60,15%	47,13%	66,53%

Tabela 3 - Resultado médio das partições baseando-se nas imagens de raio-x e utilizando a fusão em nível de decisão dos modelos selecionados

Modelo	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
VGG16	69,76%	69,88%	88,41%	41,35%	78,06%
VGG16 + MLP	75,27%	73,41%	94,47%	44,75%	82,62%
VGG16 + NB	73,38%	73,04%	91,62%	44,39%	81,28%
VGG16 + SVM	73,13%	72,33%	94,38%	38,93%	81,90%
VGG16 + RF	72,49%	71,87%	91,96%	41,72%	80,68%

ter ocorrido devido ao desbalanceamento natural dos dados, pois a classe negativa é menos representada na base.

Fusão em Nível de Características

Os resultados obtidos ao realizarmos a fusão em nível de características são sumarizados na Tabela 4. Essa tabela também mostra os resultados sem fusão. O modelo apresentou uma leve queda de desempenho quando os dados clínicos são fundidos com as características extraídas das imagens. Em geral, os resultados mostram novamente a dificuldade de classificação de instâncias negativas, observada nos valores de especificidade, provavelmente devido ao desbalanceamento do conjunto de dados.

Além disso, a rede se super-ajusta rapidamente ao conjunto de treino. Por esse motivo foi necessário utilizar tanto regularização L1 quanto L2, e taxa de aprendizagem com valor pequeno. Podemos confirmar esse comportamento através da curva da função de perda no conjunto de validação. Pela Figura 1a) podemos ver que aproximadamente após a época de treino 40, o valor da função de perda do conjunto de validação começa a subir, indicando que o modelo apresenta um leve *overfit*. Apesar disso, é possível verificar na Figura 1b) uma tendência de aumento na acurácia no conjunto de validação mesmo durante o aumento da perda.

Comparação de Resultados

Como o objetivo deste trabalho é a comparação entre

as duas abordagens de fusão, essa comparação é feita por meio das métricas estabelecidas. Os resultados obtidos pelas duas estratégias de fusão estão na tabela 5. No caso da fusão em nível de decisão, o resultado exibido é o resultado obtido pela fusão entre a CNN e um classificador MLP. Também é apresentado o melhor resultado obtido sem fusão de dados - CNN treinada com imagens de raio-x.

Essa tabela mostra que, em linhas gerais, há ganho de desempenho quando a fusão é utilizada, mais especificamente fusão em nível de decisão. Conforme observado em Zhang *et al.*⁽¹⁰⁾, nota-se que ao realizar a fusão em nível de decisão entre a CNN e os modelos de AM selecionados, consegue-se obter desempenho superior tanto em relação aos modelos individuais de CNN quanto aos classificadores clássicos de AM. Os nossos resultados seguem um padrão indicado em Pacheco *et al.*⁽⁴⁾ - aumento de 6% a 7% na acurácia média do método com fusão, se comparado ao método sem fusão com dados clínicos. Em Kharazmi *et al.*⁽⁷⁾ o mesmo comportamento se repete, havendo um aumento de 6% na acurácia utilizando fusão. Apesar desse comportamento, verifica-se claramente uma dificuldade na classificação de instâncias negativas decorrente do desbalanceamento natural do conjunto de dados. Esse é um problema complexo de ser solucionado, já que estamos trabalhando com dois tipos diferentes de dados, fato que dificulta o uso de aumento de dados.

A partir das observações dos resultados de cada método,

Tabela 4 - Resultados obtidos com fusão de dados em nível de características e sem fusão.

Modelo	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Com fusão	68,70%	68%	89,40%	37,02%	77,69%
Sem fusão	70,42%	70%	90%	41,04%	78,60%

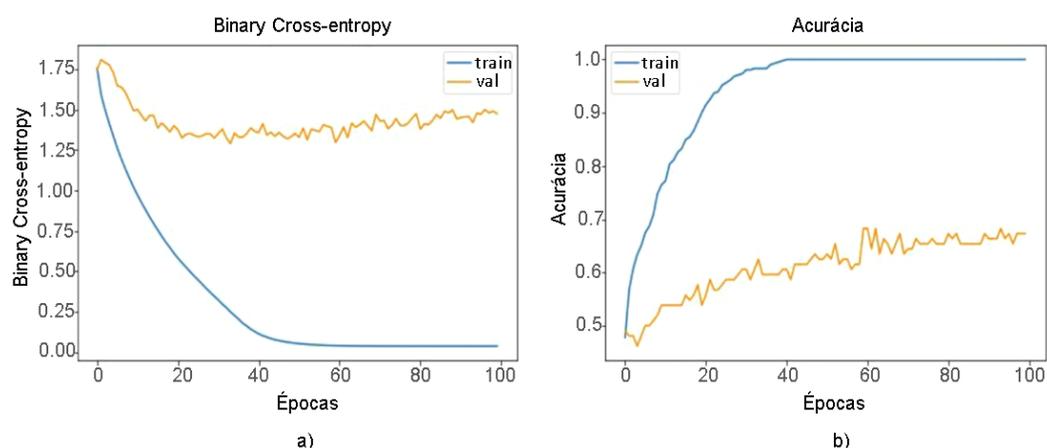


Figura 1 - a) Função de perda do método de fusão em nível de características. Em laranja a curva da função no conjunto de validação e em azul no treino. b) Acurácia do método. Em laranja a curva da acurácia no conjunto de validação e em azul no conjunto de treino.

Tabela 5 - Comparação entre o resultado dos modelos aqui apresentados.

Modelo	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Fusão em nível de decisão (VGG16+MLP)	75,27%	73,41%	94,47%	44,75%	82,62%
Fusão em nível de característica	68,70%	68%	89,40%	37,02%	77,69%
Sem fusão (VGG16)	70,42%	70%	90%	41,04%	78,60%

é possível verificar que em uma abordagem de fusão em nível de decisão os dados clínicos apresentam um peso muito maior na decisão final do modelo. Por outro lado, em fusão em nível de características, como o vetor de características extraídas de imagens e o vetor de atributos dos dados clínicos apresentam dimensões significativamente diferentes, os dados clínicos não apresentam tanta influência na decisão final. Possíveis soluções para melhorias nesses resultados seriam balanceamento entre as dimensões de cada vetor na concatenação dos dados, extração de um número maior de atributos das notas clínicas ou a redução da dimensão do vetor de características de imagens.

CONCLUSÃO

Este trabalho apresentou um estudo comparativo entre as estratégias de fusão em nível de características e fusão em nível de decisão utilizando informações de imagens de raio-x e

dados clínicos para detecção de Covid-19. Os experimentos mostraram que a estratégia de fusão em nível de decisão foi melhor do que a fusão em nível de características, especialmente porque consegue melhor compromisso entre a contribuição de cada fonte de informação para a melhoria da decisão. Além disso, a fusão melhorou os resultados dos classificadores individuais. Como trabalhos futuros, nós pretendemos combinar as duas abordagens.

AGRADECIMENTOS

Esta pesquisa, realizada no âmbito do Projeto Samsung-UFAM de Ensino e Pesquisa (SUPER), conforme previsto no Artigo 48 do Decreto nº 6.008 / 2006 (SUFRAMA), foi financiada pela Samsung Eletrônica da Amazônia Ltda., nos termos da Lei Federal nº 8.387 / 1991, através do convênio 001/2020, firmado com a Universidade Federal do Amazonas e a FAEPI, Brasil.

REFERÊNCIAS

1. Wu D, Wu T, Liu Q, Yang Z. The SARS-CoV-2 outbreak: what we know. *International Journal of Infectious Diseases*. 2020 Mar 12.
2. Pereira RM, Bertolini D, Teixeira LO, Silla Jr CN, Costa YM. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. *Computer Methods and Programs in Biomedicine*. 2020 May 8:105532.
3. Khanday AM, Rabani ST, Khan QR, Rouf N, Din MM. Machine learning based approaches for detecting COVID-19 using clinical text data. *International Journal of Information Technology*. 2020 Jun 30:1-9.
4. Pacheco AG, Krohling RA. The impact of patient clinical information on automated skin cancer detection. *Computers in biology and medicine*. 2020 Jan 1;116:103545.
5. Ahsan MM, E Alam T, Trafalis T, Huebner P. Deep MLP-CNN Model Using Mixed-Data to Distinguish between COVID-19 and Non-COVID-19 Patients. *Symmetry*. 2020 Sep;12(9):1526.
6. Mei X, Lee HC, Diao KY, Huang M, Lin B, Liu C, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nature Medicine*. 2020 May 19:1-5.
7. Kharazmi P, Kalia S, Lui H, Wang ZJ, Lee TK. A feature fusion system for basal cell carcinoma detection through data driven feature learning and patient profile. *Skin research and technology*. 2018 May;24(2):256-64.
8. Calhoun VD, Sui J. Multimodal Fusion of Brain Imaging Data: A Key to Finding the Missing Link(s) in Complex Mental Illness. *Biological psychiatry: cognitive neuroscience and neuroimaging*. 2016 May 1;1(3):230-44.
9. Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. Covid-19 image data collection: Prospective predictions are the future. *arXiv preprint arXiv:2006.11988*. 2020 Jun 22.
10. Zhang B, Qi S, Monkam P, Li C, Yang F, Yao YD, Qian W. Ensemble Learners of Multiple Deep CNNs for Pulmonary Nodules Classification Using CT Images. *IEEE Access*. 2019 Aug 7;7:110358-71.