

Classificação de Imagens Musculoesqueléticas Utilizando Aprendizado de Máquina Profundo

Leodécio Braz da Silva Segundo

Orientador: Prof. Dr. Hélio Pedrini
Coorientador: Prof. Dr. Zanoni Dias

Instituto de Computação – Unicamp

leodeciobraz@liv.ic.unicamp.br

05 de Abril de 2021

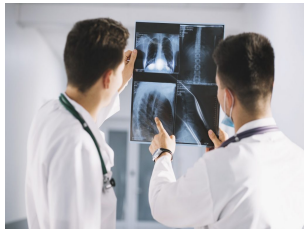
Sumário

- 1 Introdução
- 2 Trabalhos Relacionados
- 3 Conjunto de Dados e Métricas
- 4 Método Proposto
- 5 Resultados Experimentais
- 6 Conclusões e Trabalhos Futuros

Descrição do Problema



Descrição do Problema



- Sistemas de auxílio ao diagnóstico médico:
 - Processamento de Imagens.
 - Visão Computacional.
 - Aprendizado Profundo.

Descrição do Problema

- Sistemas de auxílio ao diagnóstico médico:
 - Processamento de Imagens.
 - Visão Computacional.
 - Aprendizado Profundo.
- Várias tarefas:
 - Detecção.
 - Segmentação.
 - Classificação de lesões.

Descrição do Problema

- Sistemas de auxílio ao diagnóstico médico:
 - Processamento de Imagens.
 - Visão Computacional.
 - Aprendizado Profundo.
- Várias tarefas:
 - Detecção.
 - Segmentação.
 - Classificação de lesões.
- Diversas metodologias de classificação:
 - K-Vizinhos mais Próximos.
 - Máquinas de Vetores de Suporte.
 - Florestas Aleatórias.
 - Redes Neurais.

- Métodos de Aprendizado Profundo:
 - Estado da arte.
 - Redes Neurais Convolucionais.
 - Processamento de Linguagem Natural.

Descrição do Problema

- Métodos de Aprendizado Profundo:
 - Estado da arte.
 - Redes Neurais Convolucionais.
 - Processamento de Linguagem Natural.
- Dados ou recursos limitados podem ser um problema.

Objetivos

Avaliar técnicas e processos que melhorem a detecção de anormalidades em imagens musculoesqueléticas, além de propor e implementar um método que combina informações multimodais para realizar uma classificação.

Avaliar técnicas e processos que melhorem a detecção de anormalidades em imagens musculoesqueléticas, além de propor e implementar um método que combina informações multimodais para realizar uma classificação.

- Avaliação de desempenho da classificação utilizando modelos clássicos já existentes na literatura.
- Definição de técnicas e processos a serem aplicados nas imagens.
- Avaliação de diferentes abordagens de fusão dos modelos.
- Geração de textos artificiais a partir das imagens.
- Avaliação da fusão entre métodos multimodais.

- Desenvolvimento e avaliação de diferentes métodos de fusão.
- Avaliação de um método multimodal para o problema de classificação de imagens musculoesqueléticas.

Questões de Pesquisa

Questão de Pesquisa - 1

Qual é o impacto de diferentes formas de fusão no processo de classificação?

Questão de Pesquisa - 2

A utilização de uma abordagem multimodal, no contexto de dados médicos, pode produzir resultados melhores aos resultados utilizando apenas imagens e textos?

Questão de Pesquisa - 3

É possível gerar, com boa qualidade, dados textuais de legendas artificiais para imagens de radiografias musculoesqueléticas?

Sumário

- 1 Introdução
- 2 Trabalhos Relacionados**
- 3 Conjunto de Dados e Métricas
- 4 Método Proposto
- 5 Resultados Experimentais
- 6 Conclusões e Trabalhos Futuros

Classificação de Imagens e Textos Médicos

- Estudos como os realizados por Kawahara et al. [1] e Menegola et al. [2] investigaram o problema de classificação de imagens de lesões de pele.
- Gale et al. [3] e Rajpurkar et al. [4] utilizaram imagens radiográficas de fraturas ósseas.
- Wang et al. [5] e Chen et al. [6] exploraram a classificação de doenças em imagens de raios-X de tórax.
- Smit et al. [7] propuseram um método chamado CheXbert para classificação de laudos de radiologia.

Classificação de Dados Multimodais

- Arevalo et al. [8] apresentaram o modelo chamado *Gated Multimodal Unit* (GMU).
- Kiela et al. [9] investigaram vários métodos para classificação multimodal em larga escala.
- Vielzeuf et al. [10] propuseram uma abordagem de fusão multimodal que integra informações provenientes de dados de múltiplas mídias.
- Perez-Rua et al. [11] apresentaram um método de busca para encontrar arquiteturas de fusão para classificação multimodal.
- Kiela et al. [12] propuseram um modelo multimodal que funde informações de codificadores de texto e imagem.

- Pelka et al. [13] propuseram uma abordagem para combinar palavras-chave geradas automaticamente com imagens radiográficas:
 - Método para permitir representações de imagens multimodais.
 - Fusão de informações textuais com imagens radiográficas.
 - Reconhecimento de partes do corpo e identificação anormalidades nas imagens.

Sumário

- 1 Introdução
- 2 Trabalhos Relacionados
- 3 Conjunto de Dados e Métricas**
- 4 Método Proposto
- 5 Resultados Experimentais
- 6 Conclusões e Trabalhos Futuros

Conjunto de Dados

Musculoskeletal Radiographs (MURA)

- Formado por imagens de raios-X de estudos musculoesqueléticos de extremidades superiores do corpo:
 - Cotovelo, dedo, antebraço, mão, úmero, ombro e pulso.

Conjunto de Dados

Musculoskeletal Radiographs (MURA)

- Formado por imagens de raios-X de estudos musculoesqueléticos de extremidades superiores do corpo:
 - Cotovelo, dedo, antebraço, mão, úmero, ombro e pulso.
- Cada estudo na base MURA contém uma ou mais imagens radiográficas e foram rotulados como “normais” ou “anormais” por radiologistas.

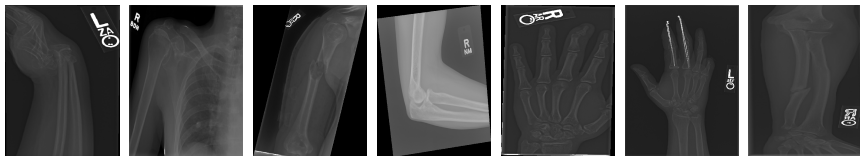
Conjunto de Dados

Musculoskeletal Radiographs (MURA)

- Formado por imagens de raios-X de estudos musculoesqueléticos de extremidades superiores do corpo:
 - Cotovelo, dedo, antebraço, mão, úmero, ombro e pulso.
- Cada estudo na base MURA contém uma ou mais imagens radiográficas e foram rotulados como “normais” ou “anormais” por radiologistas.
- A base possui 40561 imagens radiográficas musculoesqueléticas de 14863 estudos de caso a partir de 12173 pacientes:
 - Treinamento: 36808.
 - Validação: 3197.
 - Teste: 556.

Conjunto de Dados

Musculoskeletal Radiographs (MURA)



Pulso

Ombro

Úmero

Cotovelo

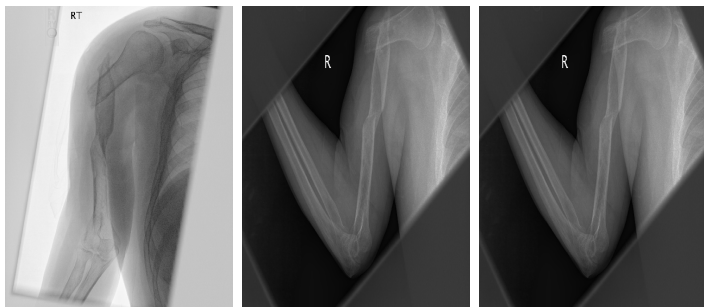
Dedo

Mão

Antebraço

Conjunto de Dados

Musculoskeletal Radiographs (MURA)



Exemplos de imagens de radiografia presentes em um mesmo estudo de caso da região do úmero rotulado como “anormal”.

Conjunto de Dados

Radiology Objects in COntext (ROCO)

- Conjunto de dados composto de imagens de radiologia oriundas de diferentes modalidades de imagens médicas:
 - Tomografia Computadorizada (CT), Ultrassonografia, Raios-X, Mamografia, Ressonância Magnética (MRI), Tomografia por Emissão de Pósitrons (PET-CT), entre outras.



Caption: Radiograph of a shoulder in lateral scapular view showing breakage of synthesis material (arrow).

Keywords: shoulder, scapular, lateral, material, radiograph, arrow, breakage, view, synthesesi

CUI: C1306645 Plain x-ray
C0524468 Structure of right shoulder region
C1548003 Radiograph
C0036277 Bone structure of scapula
C0043299 Diagnostic radiologic examination
C0037004 Shoulder

SemTypes: T060 Diagnostic Procedure
T170 Intellectual Product
T121 Pharmacologic Substance
T033 Finding
T023 Body Part, Organ, or Organ Component
T109 Organic Chemical
T029 Body Location or Region

Conjunto de Dados

CheXpert

- Grande conjunto de dados de raios-X de tórax contendo 224316 radiografias a partir de 65240 pacientes.



Cardiom.
Aumen-
tada



Cardiome-
galia



Lesão
Pulmonar



Opacidade
Pulmonar



Edema



Consolida-
ção
Pulmonar



Pneumo-
nia



Atelectasia



Pneumotó-
rax



Derrame
pleural



Outros
Prob.
Pleurais



Fratura



Dispositi-
vos de
Suporte



Nenhum
Achado

- Acurácia.
- Acurácia Balanceada.
- Área sob a curva ROC.
- Coeficiente Cohen's kappa.

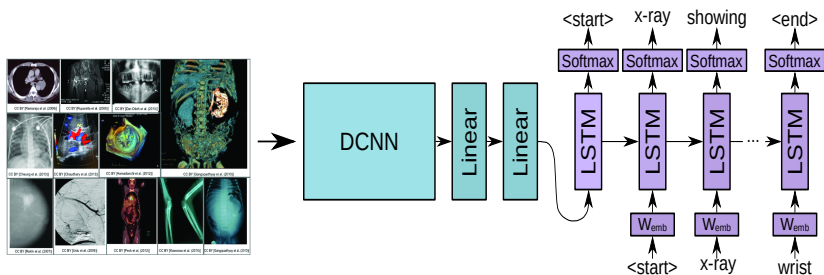
Sumário

- 1 Introdução
- 2 Trabalhos Relacionados
- 3 Conjunto de Dados e Métricas
- 4 Método Proposto**
- 5 Resultados Experimentais
- 6 Conclusões e Trabalhos Futuros

Modelos Implementados

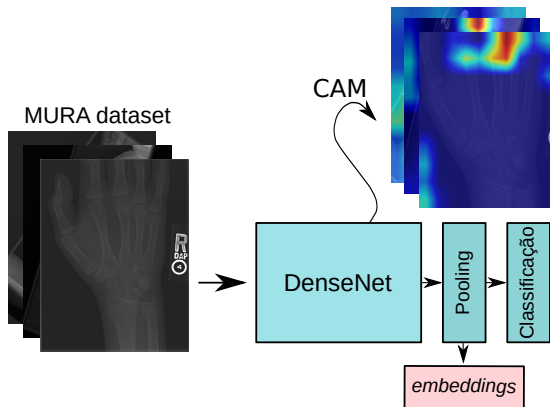
- Codificador-Decodificador.
- DenseNet.
- BERT.
- Modelo de Fusão.

Codificador-Decodificador



Representação do modelo Codificador-Decodificador para a geração de legendas a partir de imagens radiográficas.

DenseNet



Representação do modelo DenseNet utilizado para a classificação e a extração de características de imagens radiográficas.

BERT

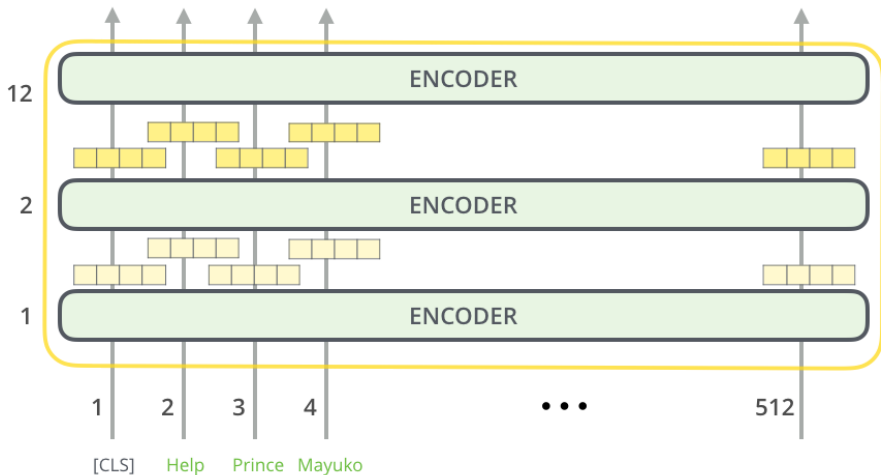
Bidirectional Encoder Representations from Transformers

Radiograph of pelvic bones showing fracture

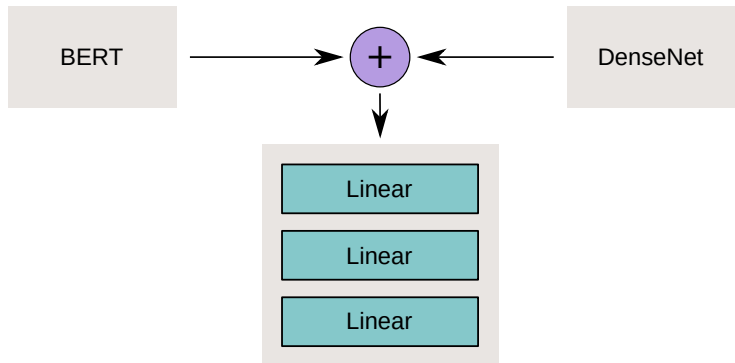
Entrada	[CLS]	radio	##graph	of	pe	##l	##vic	bones	showing	fracture	[SEP]
Token	$E_{[CLS]}$	E_{radio}	$E_{##graph}$	E_{of}	E_{pe}	$E_{##l}$	$E_{##vic}$	E_{bones}	$E_{showing}$	$E_{fracture}$	$E_{[SEP]}$
	+	+	+	+	+	+	+	+	+	+	+
Segmento	E_A	E_A	E_A	E_A	E_A	E_A	E_A	E_A	E_A	E_A	E_A
	+	+	+	+	+	+	+	+	+	+	+
codificação posicional	E_0	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}

BERT

Bidirectional Encoder Representations from Transformers



Modelo de Fusão



Geração Automática de Legendas

A person riding a motorcycle on a dirt road.



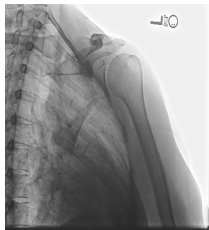
Exemplo de legendagem em imagem [14]



Exemplo de legendagem em imagem médica [15]

No acute cardiopulmonary abnormality. Stable appearance of the thoracic aorta. The right lateral lower lobe is noted in the right lower right midlung. No large pleural effusion or focal airspace disease. Mild interstitial opacities. Atherosclerotic calcifications bony structures bilaterally. There is no pleural effusion or pneumothorax developed in the right lower lobe.

Geração Automática de Legendas



Conjunto: Ombro
Classe: Normal
Legenda: the right shoulder is normal



Conjunto: Pulso
Classe: Normal
Legenda: X-ray showing the of the right wrist



Conjunto: Ombro
Classe: Anormal
Legenda: X-ray of the shoulder showing the acromioclavicular joint



Conjunto: Úmero
Classe: Anormal
Legenda: X-ray showing the fracture of the right humerus

Sumário

- 1 Introdução
- 2 Trabalhos Relacionados
- 3 Conjunto de Dados e Métricas
- 4 Método Proposto
- 5 Resultados Experimentais**
- 6 Conclusões e Trabalhos Futuros

Experimentos Utilizando Redes Neurais Convolucionais

Experimento I: Analisando Transformações

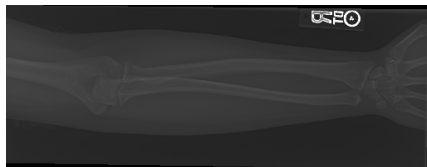


Imagem Original



Fit



Pad



Stretch

Experimentos Utilizando Redes Neurais Convolucionais

Experimento I: Analisando Transformações

	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
Fit	0,8232	0,8140	0,8720	0,6373
Pad	0,8165	0,8054	0,8723	0,6222
Stretch	0,8274	0,8175	0,8617	0,6453

Experimentos Utilizando Redes Neurais Convolucionais

Experimento II: Utilizando Redes Pré-Treinadas

Modelos	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
DenseNet-121	0,8365	0,8279	0,8874	0,6649
DenseNet-161	0,8374	0,8293	0,8892	0,6671
EfficientNet-B7	0,8274	0,8182	0,8873	0,6458
Inception-v3	0,8290	0,8197	0,8769	0,6491
Inception-v4	0,8315	0,8226	0,8783	0,6546
Inception-ResNet-v2	0,8324	0,8229	0,8864	0,6559
ResNet-18	0,8274	0,8175	0,8617	0,6453
ResNet-152	0,8299	0,8220	0,8780	0,6519
VGG-16	0,8357	0,8254	0,8877	0,6621
VGG-19	0,8282	0,8155	0,8840	0,6450

Experimentos Utilizando Redes Neurais Convolucionais

Experimento III: Transferência de Aprendizado a Partir de Outros Dados Médicos

Pré-treinamento	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
ImageNet	0,8374	0,8293	0,8892	0,6671
ImageNet+CheXpert	0,8332	0,8235	0,8906	0,6574

Experimentos Utilizando Redes Neurais Convolucionais

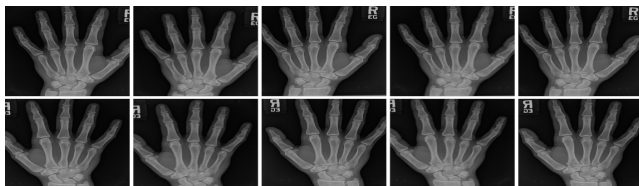
Experimento IV: Aumentação de Dados



Imagem original.



Inversão horizontal da imagem original.



Na primeira linha, os recortes realizados na imagem original.
Na segunda linha, as inversões horizontais destes.

Experimentos Utilizando Redes Neurais Convolucionais

Experimento IV: Aumentação de Dados

Base	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
Original	0,8374	0,8293	0,8892	0,6671
Aumentação A	0,8465	0,8370	0,8929	0,6848
Aumentação B	0,8432	0,8355	0,8922	0,6792

Experimentos Utilizando Redes Neurais Convolucionais

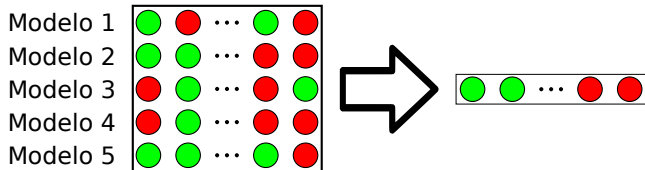
Experimento IV: Aumentação de Dados

Modelo	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
DenseNet-161	0,8465	0,8370	0,8929	0,6848
EfficientNet-B7	0,8399	0,8314	0,8913	0,6719
Inception-ResNet-v2	0,8457	0,8369	0,8892	0,6836
VGG-16	0,8374	0,8302	0,8967	0,6676
ResNet-152	0,8349	0,8243	0,8931	0,6602

Experimentos Utilizando Redes Neurais Convolucionais

Experimento V: Modelos de *Ensemble*

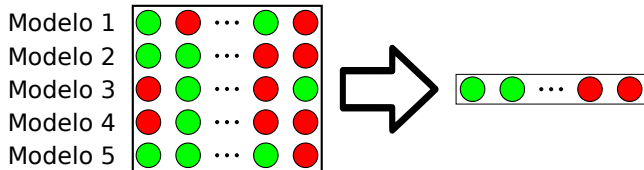
- Baseado em consenso:



Experimentos Utilizando Redes Neurais Convolucionais

Experimento V: Modelos de *Ensemble*

- Baseado em consenso:

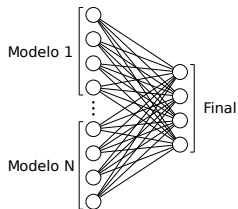


- Média ponderada das probabilidades.
 - Busca por força bruta.

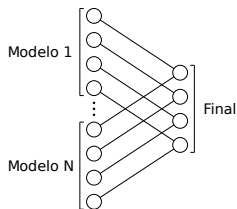
Experimentos Utilizando Redes Neurais Convolucionais

Experimento V: Modelos de *Ensemble*

- Baseado em rede neural.



Arquitetura de rede neural totalmente conectada

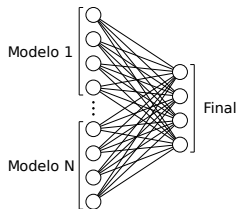


Arquitetura de rede neural com conectividade esparsa

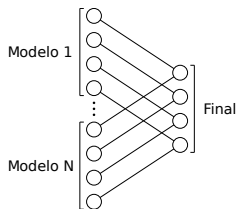
Experimentos Utilizando Redes Neurais Convolucionais

Experimento V: Modelos de *Ensemble*

- Baseado em rede neural.



Arquitetura de rede neural totalmente conectada



Arquitetura de rede neural com conectividade esparsa

- Máquinas de Vetores de Suporte (SVM).

Experimentos Utilizando Redes Neurais Convolucionais

Experimento V: Modelos de *Ensemble*

Modelo	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
Modelo Individual (média)	0,8409	0,8319	0,8926	0,6736
Modelo Individual (máximo)	0,8465	0,8370	0,8929	0,6848
Consenso	0,8449	0,8349	0,8753	0,6811
Média Ponderada	0,8641	0,8572	0,9086	0,7222
Esparsamente Conectada	0,8590	0,8504	0,9039	0,7110
Totalmente Conectada	0,8540	0,8467	0,9102	0,7015
SVM	0,8699	0,8638	0,9208	0,7345

Experimentos Utilizando Redes Neurais Convolucionais

Resultados

	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
SVM	0,8484	0,8319	0,8791	0,6724

Eficácia do modelo de *ensemble* SVM no conjunto de teste.

Experimentos Utilizando Redes Neurais Convolucionais

Resultados

	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
SVM	0,8484	0,8319	0,8791	0,6724

Eficácia do modelo de *ensemble* SVM no conjunto de teste.

	Acurácia	Acurácia Balanceada	AUC ROC	Kappa
Consenso	0,8593	0,8339	0,8652	0,6899

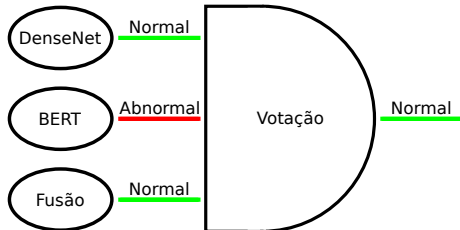
Eficácia do modelo de *ensemble* baseado em consenso no conjunto de teste.

Método Proposto

- Avaliação do método proposto:
 - DenseNet, BERT e Fusão.
- Cada estudo de caso pode possuir uma ou mais amostras de imagens:
 - Probabilidade de anormalidade em cada amostra.

Método Proposto

- Avaliação do método proposto:
 - DenseNet, BERT e Fusão.
- Cada estudo de caso pode possuir uma ou mais amostras de imagens:
 - Probabilidade de anormalidade em cada amostra.
- Esquema de votação:



Método Proposto

	DenseNet		BERT		Fusão		Método	
	Acurácia Balanceada	Kappa	Acurácia Balanceada	Kappa	Acurácia Balanceada	Kappa	Acurácia Balanceada	Kappa
Cotovelo	0,8537	0,7214	0,6322	0,2562	0,8731	0,7512	<u>0,8634</u>	<u>0,7364</u>
Dedo	0,8222	0,6521	0,6747	0,3446	0,8300	0,6651	<u>0,8294</u>	<u>0,6647</u>
Antebraço	0,8292	0,6652	0,5354	0,0705	<u>0,8471</u>	<u>0,6974</u>	0,8526	0,7114
Mão	0,7680	0,5722	0,4578	-0,0882	0,7835	0,5935	<u>0,7707</u>	<u>0,5745</u>
Úmero	0,9038	0,8074	0,5868	0,1730	<u>0,9039</u>	<u>0,8075</u>	0,9113	0,8222
Ombro	0,7724	0,5456	0,5473	0,0934	<u>0,7737</u>	<u>0,5467</u>	0,7840	0,5673
Pulso	0,8584	0,7390	0,5997	0,1973	0,8806	0,7772	<u>0,8738</u>	<u>0,7672</u>
Média	0,8296	0,6718	0,5762	0,1495	0,8417	<u>0,6912</u>	<u>0,8407</u>	0,6920

Eficácia dos modelos DenseNet, BERT, Fusão e do método proposto, que aplica uma votação entre os três classificadores.

Sumário

- 1 Introdução
- 2 Trabalhos Relacionados
- 3 Conjunto de Dados e Métricas
- 4 Método Proposto
- 5 Resultados Experimentais
- 6 Conclusões e Trabalhos Futuros**

- Abordagens para a classificação de imagens musculoesqueléticas.
- Redes Neurais Convolucionais:
 - Técnicas clássicas.
 - Etapas de pré-processamento.
 - Experimentações de forma incremental.
- Abordagem multimodal:
 - Desafio de gerar legendas.
 - Combinação de imagens com textos gerados automaticamente.
 - Ponto positivos e negativos.

Questões de Pesquisa

Questão de Pesquisa - 1

Qual é o impacto de diferentes formas de fusão no processo de classificação?

O *ensemble* de classificadores utilizando SVM apresentou resultados promissores durante os experimentos realizados, entretanto, esta técnica possui pontos negativos relacionados ao alto sobreajuste do modelo.

Questões de Pesquisa

Questão de Pesquisa - 2

A utilização de uma abordagem multimodal, no contexto de dados médicos, pode produzir resultados melhores aos resultados utilizando apenas imagens e textos?

A utilização de uma abordagem multimodal apresentou um ganho nos valores das métricas analisadas em comparação à eficácia das abordagens individuais.

Questões de Pesquisa

Questão de Pesquisa - 3

É possível gerar, com boa qualidade, dados textuais de legendas artificiais para imagens de radiografias musculoesqueléticas?

Muitos trabalhos demonstraram bons resultados para a tarefa de legendagem de imagens em geral. Entretanto, ao analisar alguns trabalhos que abordam esta tarefa no contexto de imagens médicas, constatamos que há uma carência de conjuntos de dados anotados, com os mais variados tipos de dados médicos.

- Abordagem utilizando redes neurais convolucionais:
 - Prevenir o sobreajuste.
 - Analisar outros conjuntos de dados médicos.

- Abordagem utilizando redes neurais convolucionais:
 - Prevenir o sobreajuste.
 - Analisar outros conjuntos de dados médicos.
- Abordagem multimodal:
 - Geração de legendas.
 - Modelos baseados em *transformers*.
 - Explorar outros métodos de fusão.

- Abordagem utilizando redes neurais convolucionais:
 - Prevenir o sobreajuste.
 - Analisar outros conjuntos de dados médicos.
- Abordagem multimodal:
 - Geração de legendas.
 - Modelos baseados em *transformers*.
 - Explorar outros métodos de fusão.
- Potencial para ser estendido para outras especialidades médicas, combinando os metadados com outras modalidades existentes de imagens.

- V. Teixeira, **L. Braz**, H. Pedrini, Z. Dias. DuaLANet: Dual Lesion Attention Network for Thoracic Disease Classification in Chest X-Rays. 27th International Conference on Systems, Signals and Image Processing (IWSSIP). Rio de Janeiro-RJ, Brazil, pp. 69-74, June 3-5, 2020.
- G. Sato, **L. Braz**, Z. Dias. Classification of Musculoskeletal Abnormalities with Convolutional Neural Networks. Brazilian Symposium on Bioinformatics (BSB). Online Meeting, pp. 69-80, November 23-27, 2020.
- **L. Braz**, V. Teixeira, H. Pedrini, Z. Dias. ImTeNet: Image-Text Classification Network for Abnormality Detection and Automatic Reporting on Musculoskeletal Radiographs. Brazilian Symposium on Bioinformatics (BSB). Online Meeting, pp. 150-161, November 23-27, 2020.
- **L. Braz**, V. Teixeira, H. Pedrini, Z. Dias. Image-Text Integration Using a Multimodal Fusion Network Module for Movie Genre Classification. 11th International Conference on Pattern Recognition Systems (ICPRS). Curicó, Chile, March 17-19, 2021.

- [1] J. Kawahara, A. BenTaieb, and G. Hamarneh.
Deep features to classify skin lesions.
In 13th IEEE International Symposium on Biomedical Imaging (ISBI),
pages 1397–1400, 2016.
- [2] Afonso Menegola, Michel Fornaciali, Ramon Pires, Flávia Vasques Bittencourt, Sandra Avila, and Eduardo Valle.
Knowledge Transfer for Melanoma Screening with Deep Learning.
In 14th IEEE International Symposium on Biomedical Imaging (ISBI),
pages 297–300. IEEE, 2017.
- [3] William Gale, Luke Oakden-Rayner, Gustavo Carneiro, Andrew Bradley, and Lyle Palmer.
Detecting Hip Fractures with Radiologist-level Performance using Deep Neural Networks.
arXiv:1711.06504, 2017.

- [4] Pranav Rajpurkar, Jeremy Irvin, Aarti Bagul, Daisy Ding, Tony Duan, Hershel Mehta, Brandon Yang, Kaylie Zhu, Dillon Laird, Robyn L. Ball, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, and Andrew Y. Ng.

MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs.

arXiv preprint arXiv:1712.06957, 2017.

- [5] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M. Summers.

ChestX-ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases.

In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2097–2106, 2017.

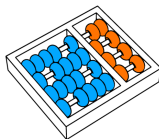
- [6] Bingzhi Chen, Jinxing Li, Xiaobao Guo, and Guangming Lu.
DualCheXNet: Dual Asymmetric Feature Learning for Thoracic Disease Classification in Chest X-rays.
Biomedical Signal Processing and Control, 53:101554, 2019.
- [7] Akshay Smit, Saahil Jain, Pranav Rajpurkar, Anuj Pareek, Andrew Ng, and Matthew Lungren.
CheXbert: Combining Automatic Labelers and Expert Annotations for Accurate Radiology Report Labeling Using BERT.
arXiv preprint arXiv:2004.09167, 2020.
- [8] John Arevalo, Tamar Solorio, Manuel Montes-y Gómez, and Fabio A González.
Gated Multimodal Units for Information Fusion.
arXiv:1702.01992, 2017.

- [9] Douwe Kiela, Edouard Grave, Armand Joulin, and Tomas Mikolov. Efficient Large-Scale Multi-Modal Classification. *arXiv:1802.02892*, 2018.
- [10] Valentin Vielzeuf, Alexis Lechervy, Stephane Pateux, and Frederic Jurie. CentralNet: a Multilayer Approach for Multimodal Fusion. In *European Conference on Computer Vision (ECCV) Workshops*, pages 1–15, 2018.
- [11] Juan-Manuel Pérez-Rúa, Valentin Vielzeuf, Stéphane Pateux, Moez Baccouche, and Frédéric Jurie. MFAS: Multimodal Fusion Architecture Search. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6966–6975, 2019.

- [12] Douwe Kiela, Suvrat Bhooshan, Hamed Firooz, and Davide Testuggine.
Supervised Multimodal Bitransformers for Classifying Images and Text.
arXiv:1909.02950, 2019.
- [13] Obioma Pelka, Felix Nensa, and Christoph M. Friedrich.
Branding - Fusion of Meta Data and Musculoskeletal Radiographs for
Multi-modal Diagnostic Recognition.
International Conference on Computer Vision Workshop (ICCV), pages
467–475, 2019.
- [14] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan.
Show and Tell: A Neural Image Caption Generator.
*IEEE Computer Society Conference on Computer Vision and Pattern
Recognition (CVPR)*, 7(12):3156–3164, June 2015.

- [15] Baoyu Jing, Pengtao Xie, and Eric P. Xing.
On the Automatic Generation of Medical Imaging Reports.
In 56th Annual Meeting of the Association for Computational Linguistics (ACL), volume 1, pages 2577–2586, 2018.

Agradecimentos



Classificação de Imagens Musculoesqueléticas Utilizando Aprendizado de Máquina Profundo

Leodécio Braz da Silva Segundo

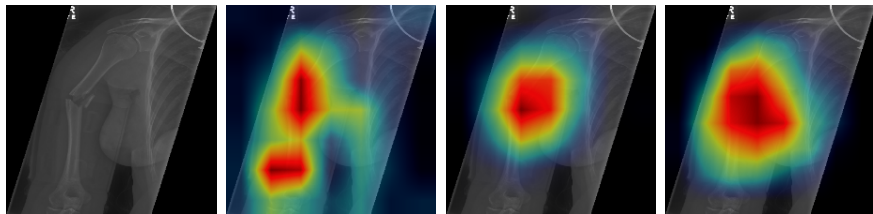
Orientador: Prof. Dr. Hélio Pedrini
Coorientador: Prof. Dr. Zanoni Dias

Instituto de Computação – Unicamp

leodeciobraz@liv.ic.unicamp.br

Experimentos Utilizando Redes Neurais Convolucionais

Mapas de Ativação dos Modelos

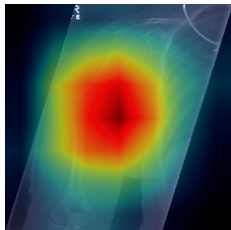


Original

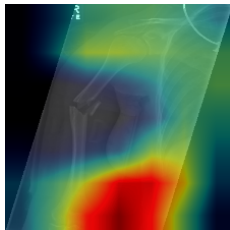
VGG-16

EfficientNet-B7

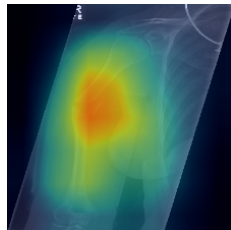
InceptionResNet-v2



DenseNet-161

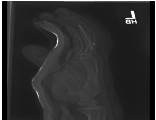
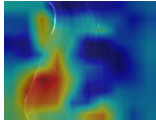

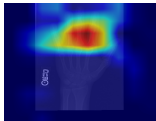
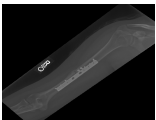
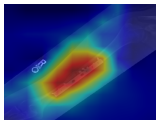

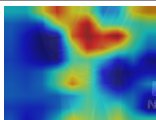


ResNet-152



Ensemble

Considerações Sobre o Método Proposto

Imagem original	Mapa de calor	Legenda gerada e classificação
		<p><start> postoperative x-ray of the patient showing the fracture of the right femur <end></p> <p>BERT: anormal ✓ DenseNet: anormal ✓ Fusão: anormal ✓ Final: anormal ✓</p>
		<p><start> the same case as in figure 1 the right hand is normal <end></p> <p>BERT: normal ✓ DenseNet: normal ✓ Fusão: normal ✓ Final: normal ✓</p>
		<p><start> the of the the of the the of the <end></p> <p>BERT: normal ✗ DenseNet: anormal ✓ Fusão: anormal ✓ Final: anormal ✓</p>
		<p><start> lateral view of the ankle showing a lytic lesion in the distal tibia<end></p> <p>BERT: anormal ✓ DenseNet: normal ✗ Fusão: normal ✗ Final: normal ✗</p>