



Departamento de Contratos e Convênios
UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO



ACORDO DE COOPERAÇÃO SEM REPASSE DE RECURSO

ACORDO DE COOPERAÇÃO Nº 15 /2019 QUE ENTRE SI
CELEBRAM A UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO E O
CENTRO UNIVERSITÁRIO DE ANÁPOLIS
PROCESSO: 23068.086644/2018-35

A **UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO**, Instituição de Ensino Superior, sob a forma de Fundação Autárquica, criada pela Lei nº. 3868, de 30/01/1961, inscrita no CNPJ/MF sob o nº. 32479123/0001-43, com sede na Av. Fernando Ferrari, nº. 514, Campus Universitário de Goiabeiras, Vitória-ES, CEP: 29075-910, ora designada como UNIVERSIDADE, doravante denominada **UFES**, neste ato representada pelo Reitor Reinaldo Centoducatte, brasileiro, casado, professor, portadora do RG nº 244.493 – SSP/ES, e do CPF nº 616.006.107-06, domiciliado no Município de Vitória-ES, e **CENTRO UNIVERSITÁRIO DE ANÁPOLIS**, doravante denominado **UniEvangélica**, com sede na Av. Universitária, km. 3, 5, Cidade Universitária, Anápolis-GO, CEP: 75083-515, neste ato representado pelo Reitor Carlos Hassel Mendes da Silva, resolvem firmar o presente instrumento, observadas as cláusulas e condições que se seguem:

CLÁUSULA PRIMEIRA – DO OBJETO

O objeto deste **ACORDO DE COOPERAÇÃO** é realização de pesquisa e desenvolvimento de tecnologia na área de engenharia biomédica. Mais especificamente, este acordo de cooperação compreende a criação de bases de dados de exames de imagem e diagnósticos associados, e o desenvolvimento de sistemas baseados em redes neurais profundas para apoio ao diagnóstico médico.

CLÁUSULA SEGUNDA – DAS COMPETÊNCIAS

SUBCLÁUSULA PRIMERA: Compete à UFES:

1. Contribuir para com a especificação dos conteúdos das bases de dados de imagens médicas digitalizadas de interesse;
2. Armazenar as bases de dados e atuar como seu curador;
3. Investigar arquiteturas de redes neurais profundas apropriadas para cada base de dados;
4. Treinar e testar as arquiteturas de redes neurais profundas desenvolvidas usando computação de alto desempenho;
5. Elaborar artigos científicos reportando os desempenhos alcançados pelas arquiteturas neurais desenvolvidas, e submetê-los a congressos e jornais científicos; e
6. Formar mestres e doutores com conhecimentos nas áreas investigadas.

SUBCLÁUSULA SEGUNDA: Compete à UniEvangélica:

1. Contribuir para com a especificação dos conteúdos das bases de dados de imagens médicas digitalizadas de interesse;
2. Coletar as bases de dados e disponibilizá-las para armazenamento e gerenciamento pela UFES;
3. Participar das investigações sobre arquiteturas de redes neurais profundas apropriadas para cada base de dados;
4. Acompanhar o treinamento e teste das arquiteturas de redes neurais profundas desenvolvidas; e
5. Participar da elaboração de artigos científicos reportando os desempenhos alcançados pelas arquiteturas neurais desenvolvidas.

CLÁUSULA TERCEIRA – DA COORDENAÇÃO

A coordenação deste instrumento, no âmbito da UFES, será de responsabilidade da Professora Claudine Santos Badue Gonçalves, CPF 819.000.661-49, SIAPE 1.729.561.

CLÁUSULA QUARTA – DA VIGÊNCIA

Este instrumento terá vigência de **3 (anos)** anos a contar da data de sua assinatura, podendo ser prorrogado caso haja necessidade de dilação do prazo, mediante termo aditivo a ser aprovado previamente pelas partes signatárias.



Handwritten signature

Handwritten signature: Badue

Handwritten signature: Gull



CLÁUSULA QUINTA – DA DENÚNCIA E RESCISÃO

Este instrumento poderá ser denunciado por qualquer uma dos Partícipes, ou rescindido mediante acordo entre as mesmas, por meio de comunicação por escrito acompanhada de memorial justificativo que produzirá efeitos após 60 (sessenta) dias, contados do recebimento pelo destinatário, fazendo-se acertos e as prestações de contas relativas às obrigações assumidas.

CLÁUSULA SEXTA – DA DIVULGAÇÃO E USO DE MARCAS

A divulgação dos atos praticados em razão deste instrumento deverá restringir-se a caráter educativo, informativo ou de disseminação da informação e conhecimento, respeitados os direitos autorais.

SUBCLÁUSULA PRIMEIRA: Os Partícipes obrigam-se a submeter previamente, por escrito, à aprovação um do outro, qualquer matéria, técnica ou científica, decorrente da execução deste instrumento, a ser eventualmente divulgada em publicações, relatórios, conclaves, propagandas, concursos e congêneres.

SUBCLÁUSULA SEGUNDA: Os Partícipes acordam que a utilização de suas respectivas marcas, representadas por seus títulos e logotipos, somente poderão ser utilizados por um partícipe com a prévia e expressa autorização do outro.

SUBCLÁUSULA TERCEIRA: Fica vedada aos Partícipes a utilização de nomes ou imagens que caracterizem promoção pessoal de autoridades ou servidores públicos, na forma prevista pelo § 1º do art. 37 da Constituição Federal.

CLÁUSULA SÉTIMA – DO SIGILO

Os Partícipes obrigam-se a manter sob o mais estrito sigilo dados e informações confidenciais trocadas, excepcionalmente, entre si ou por eles geradas na vigência deste instrumento, não podendo de qualquer forma, direta ou indiretamente, dar conhecimento a terceiros.

CLÁUSULA OITAVA – DA PROPRIEDADE INTELECTUAL

Qualquer invento, aperfeiçoamento ou inovação tecnológica, obtenção de produto ou processo resultante das ações desenvolvidas no âmbito do presente convênio terá sua exploração econômica regida por instrumento específico, assegurando-se-lhes a utilização sem ônus.

CLÁUSULA NONA – DOS RECURSOS HUMANOS

A eventual alocação de recursos humanos, por quaisquer dos Partícipes, para a execução do presente instrumento, não implicará em alteração da relação laborativa, empregatícia ou de qualquer natureza, com o órgão ou entidade de origem.

CLÁUSULA DÉCIMA – DAS ALTERAÇÕES

Durante sua vigência, este instrumento poderá ser alterado por comum acordo, no todo ou em parte, mediante celebração de Termo Aditivo, sendo vedada a inclusão posterior de metas que não tenham relação com o objeto inicialmente pactuado.

CLÁUSULA DÉCIMA PRIMEIRA – DA PUBLICAÇÃO

A UFES providenciará, sem ônus para a UniEvangélica, a publicação do extrato do presente contrato no Diário Oficial da União.

CLÁUSULA DÉCIMA SEGUNDA – DOS CASOS OMISSOS

Os casos omissos, que surgirem na execução deste instrumento, serão solucionados de comum acordo entre as partes. Não ocorrendo cumprimento das cláusulas aqui estabelecidas por parte de um dos Partícipes deverá a parte que se sentir prejudicada notificar à outra, por escrito.



Handwritten signature

Handwritten signature

Handwritten signature

Handwritten initials



Departamento de Contratos e Convênios
UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO



CLÁUSULA DÉCIMA TERCEIRA – DO FORO

É competente o Foro da Justiça Federal de Vitória, Seção Judiciária do Espírito Santo, para dirimir as questões decorrentes do presente Convênio ou de sua execução, com renúncia expressa de qualquer outro, por mais privilegiado que seja.

POR ESTAREM ASSIM JUSTOS E ACORDADOS AS PARTES DECLARAM TER LIDO E CONFERIDO O PRESENTE INSTRUMENTO, QUE FIRMAM EM 03 (TRÊS) VIAS DE IGUAL TEOR E FORMA, PARA OS MESMOS EFEITOS, E NA PRESENÇA DAS TESTEMUNHAS ABAIXO.

Vitória/ES, 09 de 08 de 2019.

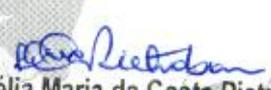

REINALDO CENTODUCATTE
Reitor da UFES


CARLOS HASSEL MENDES DA SILVA
Reitor da UniEvangélica


COORDENADOR(A) NO ÂMBITO DA UFES
Professora Claudine Santos Badue Gonçalves
SIAPE nº 1.729.561
CPF nº 819.000.661-49

TESTEMUNHAS:

NOME:
CPF:


NOME: Hêlia Maria da Costa Pietrobon
CPF: ASSESSORIA REITORIA
UniEVANGÉLICA
331.947.421-91





Departamento de Contratos e Convênios
UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO



PLANO DE TRABALHO (SEM REPASSE DE RECURSOS)

1 – DADOS CADASTRAIS DA UFES

Denominação Universidade Federal do Espírito Santo (UFES)			CNPJ 32.479.123/0001-43	
Endereço Av. Fernando Ferrari 514, Goiabeiras				
Cidade Vitória	UF ES	CEP 29060-900	Telefone (27)3335-2222	E-mail reitoria@npd.ufes.br
Responsável Institucional pela Assinatura do Convênio Reinaldo Centoducatte				
CPF 616.006.107-06			Cargo/Função Reitor	
Coordenador do Instrumento Claudine Santos Badue Gonçalves				
Cargo/Função Professor do Magistério Superior/Vice-Coodenador do LCAD			Setor de Trabalho Centro Tecnológico/Departamento de Informática/Laboratório de Computação de Alto Desempenho (LCAD)	
Matrícula 1.729.561			E-mail claudine@lcad.inf.ufes.br	
Telefone Fixo: (27)4009-2138			Celular: (27)98806-4882	

2 – DADOS CADASTRAIS DA UNIEVANGÉLICA

Denominação Centro Universitário de Anápolis (Unievangélica)			CNPJ 01.060.102/0001-65	
Endereço Av. Universitária Km 3,5, Cidade Universitária				
Cidade Anápolis	UF GO	CEP 75083-515	Telefone (62)3310-6600/3310-6711	E-mail hella@unievangelica.edu.br
Responsável Institucional pela Assinatura do Convênio Carlos Hassel Mendes da Silva				
CPF 081.124.521-72			Cargo/Função Reitor	
Coordenador do Instrumento Sérgio Mota da Silva Júnior				
Cargo/Função Professor do Magistério Superior/Vice-Coodenador da COREME			Setor de Trabalho Comissão de Residência Médica (COREME)	
Matrícula			E-mail motajunior@hotmail.com	
Telefone Fixo: (62)3310-6722			Celular: (62)98577-7535	

SR



Badue

GM

CM



3 – IDENTIFICAÇÃO DO PROJETO

3.1 – TÍTULO DO PROJETO		3.2 – PERÍODO DE EXECUÇÃO	
Análise de Imagens Médicas por meio de Cognição Visual Artificial e Interação Autônoma Inteligente		Início: A partir da data assinatura	Término: 3 (anos) anos a contar da data de sua assinatura
3.3 – OBJETO DO PROJETO			
<p>Neste projeto de pesquisa, investigaremos o uso de aprendizado profundo (<i>deep learning</i>) na análise de imagens médicas digitalizadas. Buscaremos criar uma ou mais bases de dados de imagens médicas de interesse, de pacientes saudáveis e de portadores de doenças de interesse, as quais serão descaracterizadas para preservar a privacidade dos pacientes. Para cada conjunto de imagens médicas de um paciente nas bases de dados, guardaremos também um laudo médico na forma de texto digitalizado. Guardaremos ainda dados do paciente (idade, sexo, peso, etc.), da doença e do diagnóstico médico. Buscaremos criar bases de dados com dezenas de milhares de pacientes. De posse de uma dessas bases de dados, investigaremos arquiteturas de redes neurais profundas para análise da base de dados e geração de diagnósticos de doenças.</p>			
3.4 – JUSTIFICATIVA DO PROJETO			
<p>Assim que foi possível digitalizar e carregar imagens médicas em um computador, os pesquisadores criaram sistemas para análise automatizada (Litjens et al., 2017). Inicialmente, da década de 1970 até a década de 1990, a análise de imagens médicas foi feita por sistemas baseados em regras, que resolvem tarefas particulares usando processamento de pixels de baixo nível (filtros de detecção de quinas e linhas, crescimento de região) e modelagem matemática (ajustes de linhas, círculos e elipses). No final da década de 1990, sistemas baseados em técnicas supervisionadas começaram a se tornar cada vez mais populares na análise de imagens médicas, os quais são desenvolvidos usando dados de treinamento. Exemplos de técnicas supervisionadas incluem modelos de forma ativa (para segmentação), métodos de atlas (nos quais os dados de treinamento são compostos por atlas que se ajustam a novos dados), e extração de características e classificadores estatísticos (para detecção e diagnóstico assistidos por computador). Assim, houve uma mudança de sistemas completamente projetados por seres humanos para sistemas treinados por computadores usando dados de exemplo, dos quais são extraídos vetores de características. Entretanto, nesses sistemas, a extração de características discriminantes das imagens é feito por pesquisadores humanos. Por isso, esses sistemas são chamados de sistemas com características manufaturadas.</p> <p>No final da década de 2000, começaram a aparecer sistemas baseados em aprendizado profundo (LeCun et al., 2015) para análise de grande volume de dados. No aprendizado profundo, modelos (redes) compostos por muitas camadas transformam dados de entrada (por exemplo, imagens médicas) em saídas (por exemplo, diagnósticos de doenças) enquanto aprendem características de nível cada vez mais alto. O tipo mais bem sucedido de modelos para análise de imagens são as redes neurais convolucionais (<i>Convolutional Neural Networks - CNNs</i>). As CNNs contêm muitas camadas que transformam sua entrada com filtros de convolução de pequena extensão. A primeira aplicação de CNN de sucesso no mundo real foi a LeNet (LeCun et al., 1998) para reconhecimento de dígitos escritos à mão. Apesar desses sucessos iniciais, o uso de CNNs não ganhou força até que várias novas técnicas foram desenvolvidas para o treinamento eficiente de redes neurais profundas, e avanços foram feitos nos principais sistemas de computação. O divisor de águas foi a contribuição de Krizhevsky et al. (2012) para o desafio ImageNet em 2012. A CNN proposta, chamada AlexNet, venceu essa competição por uma grande margem. Nos anos subsequentes, mais progressos foram feitos usando arquiteturas relacionadas, mas mais profundas (Russakovsky, 2014). Na visão computacional, as CNNs tornaram-se a técnica de escolha.</p> <p>Sistemas baseados em aprendizado profundo para análise de imagens médicas começaram a aparecer em 2012. Revisões sobre sistemas baseados em aprendizado profundo para análise de imagens médicas foram publicadas por Litjens et al. (2017) e Shen et al. (2017). Esses trabalhos indicam que, com o aprendizado profundo, o desempenho dos computadores na análise de imagens médicas melhorou de forma significativa e pode ser superior ou equivalente ao desempenho humano em alguns casos.</p> <p>Este projeto de pesquisa se justifica pelo potencial que ele possui de avançar o estado-da-arte na área de análise de imagens médicas usando aprendizado profundo e pelos benefícios que a sociedade pode auferir com o uso dessas técnicas, já que elas viabilizarão o aumento da produtividade do trabalho dos médicos e a melhoria da qualidade da geração de diagnósticos de doenças.</p>			



Handwritten signatures and initials: "SB", "E. Badue", "905", and "EM".

4 – DESCRIÇÃO DO PROJETO

O Laboratório de Computação de Alto Desempenho (LCAD) da UFES desenvolve pesquisa nas áreas de Cognição Visual Artificial e Interação Autônoma Inteligente, entre outras áreas. Abaixo, descrevemos sucintamente essas duas áreas de pesquisa.

Cognição Visual Artificial

A cognição pode ser definida como a nossa capacidade de compreender o mundo e as ideias por meio dos nossos sentidos e de nossa memória de experiências passadas. Podemos categorizar grosseiramente nossas diversas habilidades cognitivas de acordo com nossos sentidos, levando a categorias tais como cognição auditiva, cognição tátil ou cognição visual. Na área de pesquisa em Cognição Visual Artificial, buscamos compreender os mecanismos naturais que nos propiciam a cognição visual e investigar modelos matemático-computacionais para sua emulação. Tais modelos tipicamente recebem, como entrada, imagens estáticas ou vídeos e geram, como saída, variáveis de interesse que podem ser usadas para resolver problemas práticos relevantes.

Interação Autônoma Inteligente

Nós, humanos, somos capazes de participar de situações do dia-a-dia por meio de interações uns com os outros, e com objetos e sistemas naturais ou artificiais do mundo. Em tais interações, a participação de cada um de nós segue modelos, ou scripts (roteiros), que representamos. Para diferentes contextos, esses scripts descrevem ações apropriadas, ou sequências de ações apropriadas, que dependem das ações de nossos interlocutores, estado de objetos e/ou sistemas com os quais estamos interagindo. As ações apropriadas dependem, ainda, do ponto em que estamos na execução de um determinado script. Considere, por exemplo, o comportamento esperado de cada um de nós quando vamos a um restaurante para uma refeição. Cada etapa de nossa interação com o garçom, mesas, cadeiras, utensílios, etc., foram aprendidas ao longo de nossa vida, inclusive suas nuances e variabilidades. É esperado que sigamos um script específico quando vamos a um restaurante para uma refeição. Se vamos a um restaurante para retirar um pedido feito previamente ou fazer uma entrega (no papel de um entregador de insumos do restaurante, por exemplo), contudo, os scripts são muito diferentes, muito embora possam envolver interação com o garçom, mesas, cadeiras, etc. Na área de pesquisa em Interação Autônoma Inteligente, buscamos compreender os mecanismos naturais que nos propiciam a capacidade de realizar scripts por intermédio de nossa interação com outros por meio de linguagem natural, e com o mundo físico e seus sistemas naturais ou artificiais por meio de nosso corpo ou de hardware ou software que controlamos. Buscamos também, e fundamentalmente, investigar como implementar sistemas computacionais capazes de executar scripts envolvendo a interação com humanos e outros sistemas computacionais por meio de linguagem natural e interação física autônoma com o mundo via sistemas robóticos.

Análise de Imagens Médicas por meio de Cognição Visual Artificial e Interação Autônoma Inteligente

Dentre as diversas técnicas de análise de imagens, o aprendizado profundo, ou *deep learning* (LeCun et. al, 2015), tem ganhado proeminência recentemente. O aprendizado profundo é uma técnica de inteligência artificial baseada em modelos matemático-computacionais simplificados de redes de neurônios biológicos. Tais modelos, denominados redes neurais artificiais profundas, são constituídos de muitas camadas de neurônios artificiais empilhadas (daí a denominação "profundo") e interconectadas por sinapses simplificadas, cada uma representada por um valor ajustável (LeCun et. al, 2015). O ajuste dessas sinapses é feito por um mecanismo de aprendizado que torna o modelo capaz de reconhecer relações entre dados de entrada e dados de saída da rede neural profunda observados durante o aprendizado. Após o aprendizado, essas redes neurais podem ser usadas para, dada uma entrada, inferir uma saída de interesse associada.

Se o aprendizado ocorreu de forma apropriada, a saída gerada por uma rede neural profunda a partir de uma ou mais entradas guardará uma relação com essas entradas equivalente às relações similares entre conjuntos de entrada-saída observados durante o treinamento. Elas têm, assim, a capacidade (após serem treinadas) de extrair automaticamente relações entre dados não estruturados e rótulos associados a esses dados por especialistas. Esses dados não estruturados, que podem ser imagens, fala, texto, vídeo, etc., representam as entradas, e os rótulos – que também podem ser imagens, fala, texto, vídeo, etc. – representam as saídas mencionadas acima, os quais são usados nos conjuntos de entrada-saída durante o treinamento.

Na fase de inferência, um dado não estruturado é usado como entrada e a rede neural profunda gera uma saída, sem a intervenção humana, que guarda uma relação com o dado de entrada equivalente a uma relação similar observada durante o treinamento. Existem outros mecanismos de apresentação de dados para treinamento de redes neurais profundas que não na forma de conjuntos de entrada-saída (aprendizado não supervisionado). Um panorama geral da área de aprendizado profundo pode ser examinado em Goodfellow et al. (2016).

Recentemente, aprendizado profundo vem sendo empregado com sucesso na análise de imagens médicas. Esteva et al. (2017) empregaram uma rede neural convolucional (*Convolutional Neural Networks* - CNN) para diagnóstico de câncer de pele a partir de imagens de lesões cutâneas. A CNN foi treinada de ponta a ponta a partir de imagens diretamente, usando apenas pixels e rótulos de doenças como entradas. Treinaram a CNN usando uma base de dados de 129.450 imagens consistindo de 2.032 doenças diferentes. Testaram seu desempenho comparado ao de 21 dermatologistas certificados em biópsias de imagens com dois casos de classificação binária: carcinomas de queratinócitos versus queratoses seborreicas benignas e melanomas malignos versus nevos benignos. O primeiro caso representa a identificação dos cânceres mais comuns e o segundo caso representa a identificação do câncer de pele mais letal. A CNN alcançou desempenho equivalente ao de todos os especialistas em ambas as tarefas, demonstrando uma inteligência artificial capaz de classificar o câncer de pele com um nível de competência comparável ao de dermatologistas. A Figura 1 mostra o desempenho de classificação de câncer de pele da CNN e de dermatologistas.

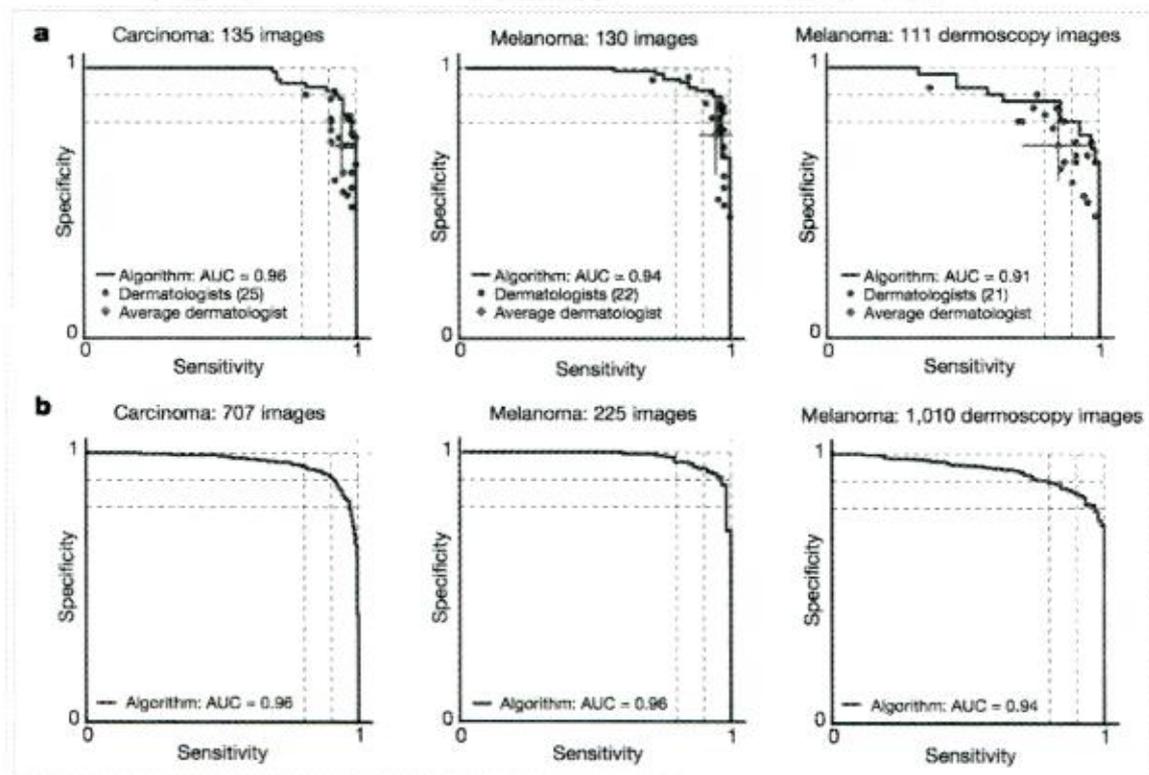


Figura 1: Desempenho de classificação de câncer de pele da CNN e dermatologistas. (a) A CNN supera a média dos dermatologistas na classificação de câncer de pele (carcinomas de queratinócitos e melanomas) usando imagens fotográficas e dermatoscópicas. Para cada teste, imagens inéditas de lesões comprovadas por biópsia são exibidas, e os dermatologistas são questionados se gostariam de: biopsiar, tratar a lesão ou tranquilizar o paciente. Um dermatologista gera uma única previsão por imagem, a qual é representada por um único ponto vermelho. Os pontos verdes são a média dos dermatologistas para cada tarefa, com barras de erro denotando um desvio padrão calculado a partir de 25, 22 e 21 dermatologistas testados para carcinoma, melanoma e melanoma sob dermatoscopia, respectivamente. A CNN é representada pela curva azul, e a AUC é a medida de desempenho da CNN, com um valor máximo de 1. A CNN alcança um desempenho superior para um dermatologista se o ponto de sensibilidade-especificidade do dermatologista estiver abaixo da curva azul – a maioria dos pontos está. b) A CNN exibe classificação confiável de câncer quando testada em uma base de dados maior. As curvas da CNN são mais suaves devido a base de dados de teste maior. [Fonte: Esteva et al. (2017)]



508

Badue

AB

EU

Suk et al. (2014) empregaram uma *Deep Boltzmann Machine* (DBM), uma rede neural profunda com uma máquina restrita de Boltzmann como unidade básica, para diagnóstico da doença de Alzheimer (*Alzheimer's Disease* - AD) e seu estágio prodômico, o Comprometimento Cognitivo Leve (*Mild Cognitive Impairment* - MCI), a partir da fusão de informações multimodais de imagens de ressonância magnética (*Magnetic Resonance Imaging* - MRI) e tomografia por emissão de pósitrons (*Positron Emission Tomography* - PET). Especificamente, usaram uma DBM para encontrar uma representação semântica latente de características a partir de um fragmento 3D e, em seguida, projetar um método sistemático para uma representação conjunta de características a partir dos fragmentos emparelhados de MRI e PET com uma DBM multimodal. Para validar a eficácia do método proposto, realizaram experimentos com a base de dados *Alzheimer's Disease Neuroimaging Initiative* (ADNI, <http://adni.loni.usc.edu/>) e compararam com métodos do estado da arte. Em três problemas de classificação binária de AD versus controle normal saudável (*Normal Control* - NC), MCI versus NC e conversor MCI versus MCI não-conversor, obtiveram precisões máximas de 95.35%, 85.67% e 74.58%, respectivamente, superando o desempenho dos métodos concorrentes.

Wang et al. (2016) empregaram modelos baseados em aprendizado profundo para diagnóstico de câncer de mama com microcalcificações a partir de imagens de mamografia. Um método semi-automático de segmentação foi usado para caracterizar todas as microcalcificações. Um modelo classificador de discriminação foi construído para avaliar as microcalcificações e as massas mamárias, isoladamente ou em combinação, a fim de classificar as lesões mamárias. O modelo de aprendizado profundo alcançou uma acurácia discriminativa de 87.3% quando as microcalcificações foram caracterizadas isoladamente. A acurácia foi de 61.3% com massa isolada e melhorou para 89.7% após análise combinada com microcalcificações. No geral, o aprendizado profundo baseado em grandes bases de dados foi superior aos métodos padrão (*Support Vector Machines* - SVM, k-Nearest Neighbor - kNN e Linear Discriminant Analysis - LDA) para a discriminação de microcalcificações.

Esses resultados mostram que aprendizado profundo já é uma poderosa ferramenta de auxílio ao diagnóstico de diversas doenças. Resultados recentes em outras áreas de análise de imagens mostraram, ainda, que aprendizado profundo pode gerar uma descrição textual a partir de uma imagem com base nos objetos e ações da imagem. Xu et al. (2015) propuseram um modelo baseado em atenção que aprende automaticamente a descrever o conteúdo das imagens. O modelo usa uma CNN para extrair um conjunto de vetores de característica, cada um dos quais sendo uma representação multidimensional correspondendo a uma parte da imagem, e uma rede de memória de curto prazo longa (*Long Short-Term Memory* - LSTM) (Hochreiter and Schmidhuber, 1997) para produzir uma legenda ao gerar uma palavra a cada passo de tempo condicionada a um vetor de contexto, o estado oculto anterior e as palavras geradas anteriormente. Figura 2 mostra exemplos de descrições de conteúdos de imagens pelo modelo baseado em atenção. Avaliaram quantitativamente o modelo baseado em atenção em três bases de dados - Flickr8k (Hodosh et al., 2013), Flickr30k (Young et al., 2014) e Microsoft COCO (Lin et al., 2014) - e obtiveram desempenho equivalente ao do estado-da-arte nas três bases de dados.

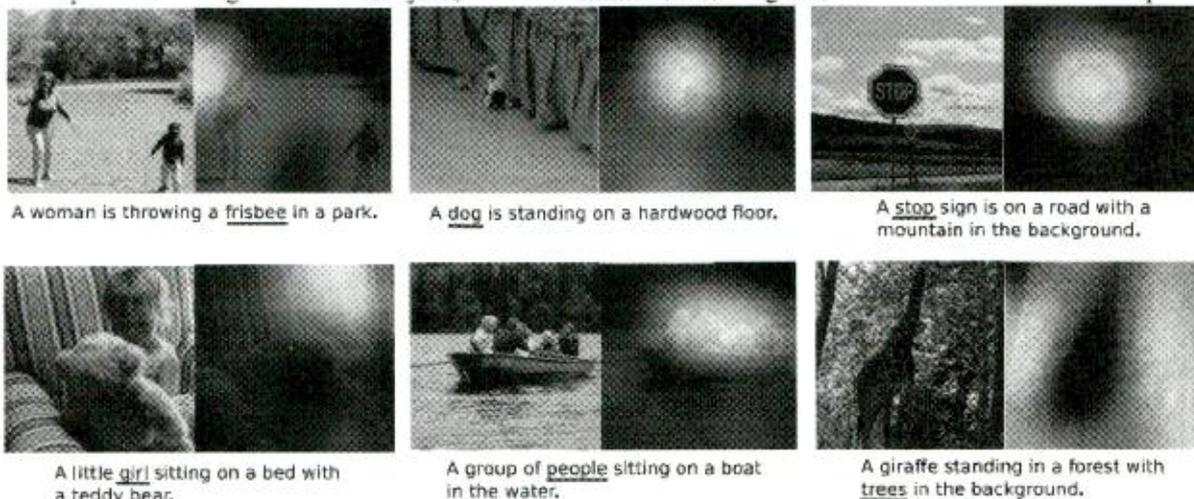


Figura 2: Exemplos de descrições de conteúdos de imagens pelo modelo baseado em atenção. [Fonte: Xu et al. (2015)]

You et al. (2016) propuseram uma abordagem para legendagem de imagens que combina abordagens *top-down* e *bottom-up* por meio de um modelo de atenção semântica. Abordagens *top-down* partem da essência de uma imagem e a converte

em palavras e abordagens *bottom-up* produzem palavras que descrevem vários aspectos de uma imagem e depois as combinam. O modelo de atenção semântica aprende a atender seletivamente a propostas de conceitos semânticos (ou atributos) e a fundi-las em estados ocultos e saídas de redes neurais recorrentes (*Recurrent Neural Networks – RNNs*). A seleção e a fusão formam um retorno que conecta a computação *top-down* e *bottom-up*. Avaliaram o modelo de atenção semântica em duas bases de dados: Microsoft COCO (Chen et al., 2015) e Flickr30K (Plummer et al., 2015). Figura 3 mostra exemplos de legendagem de imagens pelo modelo de atenção semântica (ATT) e pelo Google NIC (Vinyals et al., 2015). Resultados experimentais mostram que seu algoritmo supera abordagens do estado da arte em termos de diferentes métricas de avaliação.

Google NIC	a white plate topped with a variety of food	a baby is eating a piece of paper	a close up of a plate of food on a table.	a teddy bear sitting on top of a chair.	a person is holding colorful umbrella.	a woman is holding a cell phone in her hand	a traffic light is on a city street	a yellow and black train on a track.
Top-3 visual attributes	plate broccoli fries food french	teeth brushing toothbrush holding baby	cake table plate sitting birthday	teddy cat bear stuffed white	umbrella beach water sitting boat	woman bathroom her scissors man	street sign cars clock tower	train tracks clock tower down
ATT-FCN	a plate with a sandwich and french fries.	a baby with a toothbrush in its mouth.	a table topped with a cake with candles on it.	a white teddy bear sitting next to a stuffed animal.	a black umbrella sitting on top of a sandy beach.	a woman holding a pair of scissors in her hands.	a street with cars and a clock tower.	a train traveling down tracks next to a building.

Figura 3: Exemplos de legendagens de imagens pelo modelo de atenção semântica (ATT-FCN) e pelo Google NIC. [Fonte: You et al. (2016)]

Os resultados de Xu et al. (2015) e You et al. (2016) estão em linha com nossa área de pesquisa em Interação Autônoma Inteligente; ele sugere que redes neurais profundas podem realizar o script de elaborar um laudo de exame a partir de uma imagem médica.

Neste projeto de pesquisa, investigaremos o uso de aprendizado profundo na análise de imagens médicas digitalizadas. Buscaremos criar uma ou mais bases de dados de imagens médicas de interesse, de pacientes saudáveis e de portadores de doenças de interesse, as quais serão descaracterizadas para preservar a privacidade dos pacientes. Para cada conjunto de imagens médicas de um paciente nas bases de dados, guardaremos também um laudo médico na forma de texto digitalizado. Guardaremos ainda dados do paciente (idade, sexo, peso, etc.), da doença e do diagnóstico médico. Buscaremos criar bases de dados com dezenas de milhares de pacientes. De posse de uma dessas bases de dados, investigaremos arquiteturas de redes neurais profundas para análise da base de dados e geração de diagnósticos de doenças. Inicialmente, focaremos puramente na geração de diagnósticos e, posteriormente, de laudos médicos completos.

5 – ATRIBUIÇÕES DAS PARTES

5.1 – ATRIBUIÇÕES DA UFES

Competirá à UFES as seguintes atribuições:

1. Contribuir com a especificação dos conteúdos das bases de dados de imagens médicas de interesse digitalizadas;
2. Armazenar as bases de dados digitalizadas e atuar como seu curador;
3. Investigar arquiteturas de redes neurais profundas apropriadas para cada base de dados;
4. Treinar e testar as arquiteturas neurais desenvolvidas usando computação de alto desempenho;
5. Elaborar artigos científicos reportando os desempenhos alcançados pelas arquiteturas neurais desenvolvidas, e submetê-los a congressos e jornais científicos; e
6. Formar mestres e doutores com conhecimentos nas áreas investigadas.

5.2 – ATRIBUIÇÕES DA UniEvangélica

Competirá à UniEvangélica as seguintes atribuições:

1. Contribuir com a especificação dos conteúdos das bases de dados de imagens médicas de interesse digitalizadas;



2. Coletar as bases de dados digitalizadas, e disponibilizá-las para armazenamento e gerenciamento pela UFES;
3. Participar das investigações sobre arquiteturas de redes neurais profundas apropriadas para cada base de dados;
4. Acompanhar o treinamento e teste das arquiteturas neurais desenvolvidas; e
5. Participar da elaboração de artigos científicos reportando os desempenhos alcançados pelas arquiteturas neurais desenvolvidas.

6 – CRONOGRAMA DE EXECUÇÃO

Meta	Etapa/Fase	Especificação	Indicador Físico		Duração	
			Unidade	Quantidade	Início	Término
Criação das bases de dados de imagens médicas de interesse digitalizadas	Definição das bases de dados de interesse	Escolha das doenças e imagens médicas associadas	Número de doenças de interesse	2	Fev/2019	Mar/2019
	Especificação dos conteúdos das bases de dados	Descrição detalhada dos campos das bases de dados	Número de bases de dados especificadas	2	Abr/2019	Jul/2019
	Coleta das bases de dados	Definição e realização de procedimentos para a coleta das bases de dados	Número de bases de dados coletadas	2	Ago/2019	Jul/2020
Investigação de arquiteturas de redes neurais profundas apropriadas para cada base de dados	Exame da bibliografia correlata	Identificação e estudo de artigos científicos relevantes	Não se aplica	-	Abr/2019	Jul/2020
	Implementação de arquiteturas de redes neurais profundas, e treinamento e teste das arquiteturas neurais implementadas em pequena escala	Implementação, na linguagem de programação apropriada, de arquiteturas neurais, e treinamento e teste das arquiteturas implementadas com pequenas bases de dados	Não se aplica	-	Out/2019	Jul/2020
Treinamento e teste das arquiteturas de redes neurais profundas desenvolvidas	Escolha das arquiteturas de redes neurais profundas apropriadas para as bases de dados	Avaliação do desempenho das arquiteturas neurais desenvolvidas e identificação das arquiteturas de melhor desempenho	Número de arquiteturas escolhidas	2	Ago/2020	Set/2020
	Treinamento e teste em larga escala das arquiteturas neurais escolhidas	Treinamento e teste em larga escala das arquiteturas neurais de melhor desempenho	Número de arquiteturas treinadas e testadas	2	Out/2020	Set/2021
Publicação dos resultados deste projeto de pesquisa	Publicação dos resultados da pesquisa	Identificação de congressos e jornais científicos relevantes, redação de artigos científicos reportando os desempenhos alcançados pelas arquiteturas neurais desenvolvidas, e submissão dos artigos	Número de artigos científicos publicados	3	Ago/2020	Fev/2022





		redigidos aos congressos e jornais identificados				
--	--	--	--	--	--	--

7 – PLANO DE APLICAÇÃO

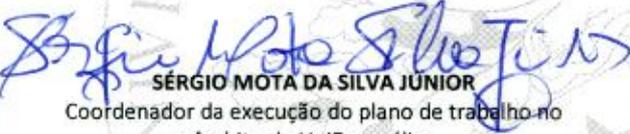
NÃO HAVERÁ TRANSFERÊNCIA DE RECURSOS ENTRE OS PARTÍCIPES.

8 – APROVAÇÃO

Vitória (ES), 09 de Agosto de 2019.


CARLOS HASSEL MENDES DA SILVA
Reitor da UniEvangélica


REINALDO CENTODUCATTE
Reitor da UFES


SÉRGIO MOTA DA SILVA JUNIOR
Coordenador da execução do plano de trabalho no âmbito da UniEvangélica


CLÁUDINE SANTOS BADUE GONÇALVES
Coordenador da execução do plano de trabalho no âmbito da UFES

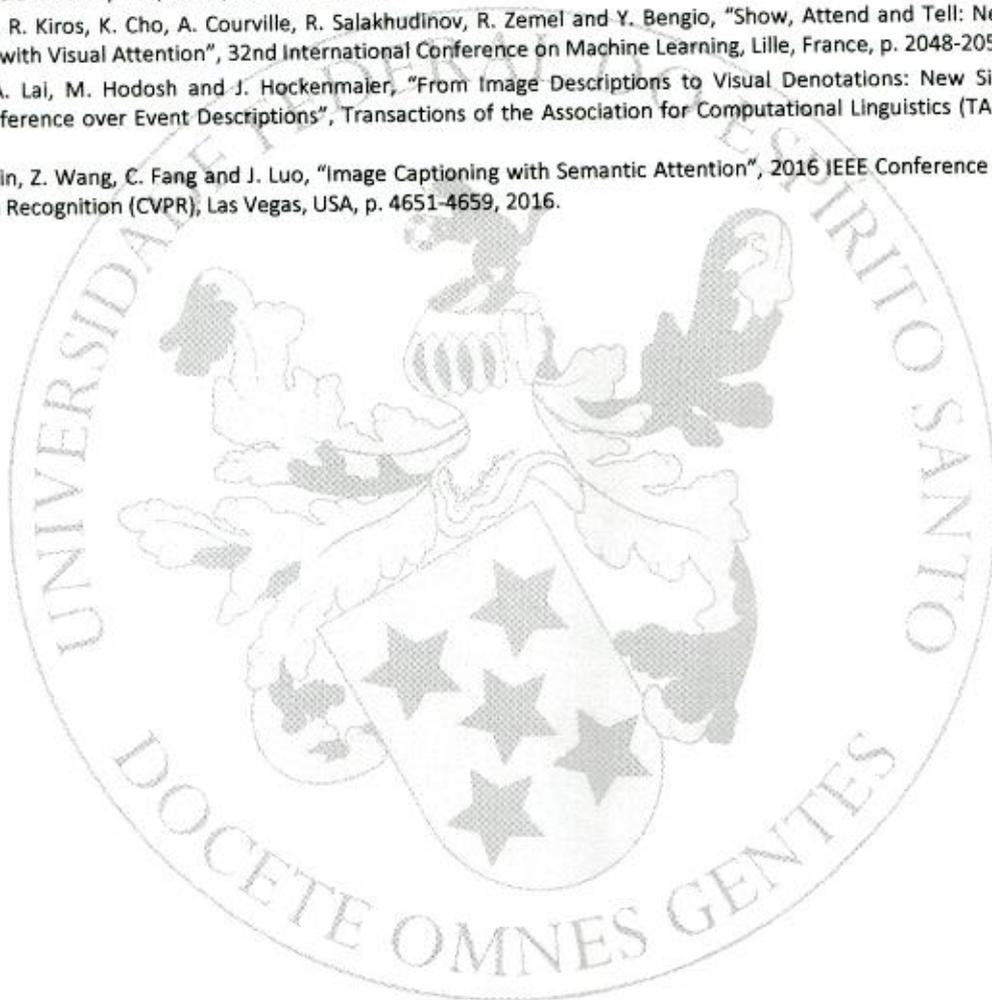
REFERÊNCIAS BIBLIOGRÁFICAS

1. X. Chen, H. Fang, T.-Y. Lin, R. Vedantam, S. Gupta, P. Dollar and C. L. Zitnick, "Microsoft COCO Captions: Data Collection and Evaluation Server", arXiv preprint arXiv:1504.00325, 2015.
2. A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau and S. Thrun, "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks", Nature, vol. 542, p. 115–118, 2017.
3. I. Goodfellow, Y. Bengio and A. Courville, "Deep Learning", Cambridge: MIT Press, 2016.
4. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", Neural Computation, vol. 9, no. 8, p. 1735–1780, 1997.
5. M. Hodosh, P. Young and J. Hockenmaier, "Framing Image Description as a Ranking Task: Data, Models and Evaluation Metrics", Journal of Artificial Intelligence Research, vol. 47, p. 853–899, 2013.
6. A. Krizhevsky, I. Sutskever and G. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems, p. 1097–1105, 2012.
7. Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning", Nature, vol. 521, no. 7553, p. 436, 2015.
8. Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-Based Learning Applied to Document Recognition", Proceedings of the IEEE, vol. 86, no. 11, p. 2278–2324, 1998.
9. T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, C. L. Zitnick, "Microsoft COCO: Common Objects in Context", European Conference on Computer Vision (ECCV), Zurich, Switzerland, p. 740–755, 2014.
10. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A.W.M. van der Laak, B. van Ginneken and C. I. Sanchez, "A Survey on Deep Learning in Medical Image Analysis", arXiv preprint arXiv:1702.05747v2, 2017.
11. B. A. Plummer, L. Wang, C. M. Cervantes, J. C. Calcedo, J. Hockenmaier and S. Lazebnik, "Flickr30k Entities: Collecting Region-to-Phrase Correspondences for Richer Image-to-Sentence Models", 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 2015.





12. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, A. C. and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge", International Journal on Computer Vision, vol. 115, no. 3, p. 1-42, 2014.
13. D. Shen, G. Wu and H.-i. Suk, "Deep Learning in Medical Image Analysis", Annual Review of Biomedical Engineering, vol. 19, p. 221-248, 2017.
14. H.-i. Suk, S.-W. Lee and D. Shen, Hierarchical Feature Representation and Multimodal Fusion with Deep Learning for AD/MCI Diagnosis, NeuroImage, vol. 101, p. 569-582, 2014.
15. O. Vinyals, A. Toshev, S. Bengio and D. Erhan, "Show and Tell: A Neural Image Caption Generator", 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, USA, p. 3156-3164, 2015.
16. J. Wang, X. Yang, H. Cai, W. Tan, C. Jin and L. Li, "Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning", Scientific Reports, vol. 6, no. 27327, 2016.
17. K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel and Y. Bengio, "Show, Attend and Tell: Neural Image Caption Generation with Visual Attention", 32nd International Conference on Machine Learning, Lille, France, p. 2048-2057, 2015.
18. P. Young, A. Lai, M. Hodosh and J. Hockenmaier, "From Image Descriptions to Visual Denotations: New Similarity Metrics for Semantic Inference over Event Descriptions", Transactions of the Association for Computational Linguistics (TACL), vol 2, p. 67-78, 2014.
19. Q. You, H. Jin, Z. Wang, C. Fang and J. Luo, "Image Captioning with Semantic Attention", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, p. 4651-4659, 2016.



46



Reitoria

Handwritten initials and a signature in the bottom right corner.