

Redes Neurais Artificiais Aplicadas para Tradução

A Inteligência Artificial utilizada nos resultados do Google Tradutor e Estruturas Audiovisuais

Diego Falabella Donada Aguiar

Cinema de Animação e Artes Digitais

Universidade Federal de Minas Gerais, UFMG, Escola de Belas Artes, Belo Horizonte, Brasil

diegoconec@gmail.com

Resumo— O trabalho a seguir percorre uma análise sobre o uso de redes neurais artificiais no aplicativo multiplataforma Google Tradutor e investiga essas mesmas redes usadas em obras de arte digital visuais e jogos digitais descrevendo-as. Foram usadas biografias técnicas do campo da Ciência da Computação ligadas à Inteligência Artificial e o uso dos aplicativos por meio da *internet* para esse exame. O trabalho conclui que as redes neurais artificiais são uma tecnologia favorável à otimização de tarefas humanas e pode ser convergida com outras tecnologias, como a realidade aumentada, para atender às diversas demandas. O trabalho também aborda que o campo de aplicação das redes neurais artificiais na arte digital é promissor e pouco desbravado, tendo amplo espaço para produção, discussão e crítica artística.

Palavras-Chaves—Redes Neurais Artificiais; Inteligência Artificial; Aprendizado de Máquina; Deep Learning; Tradução; Google Tradutor;

I. INTRODUÇÃO

Com o despertar das pesquisas na área de Inteligência Artificial, surgiram demandas e possibilidades na automação de atividades de trabalho não imaginadas até o início do século XXI. Um exemplo disso é o hábito de consultar a um dicionário ao se deparar com uma frase ou palavra de língua estrangeira não conhecida para revelar seu significado. O aplicativo para *smartphone* Google Tradutor utiliza de redes neurais artificiais¹, também denominadas apenas como “redes neurais”, para promover tradução instantânea de textos contidos em ambiente externo ao digital com interface simples de usar. Este trabalho procura analisar de que maneira as redes neurais artificiais são utilizadas para promover a tradução de textos contidos em imagens captadas pelo aplicativo Google Tradutor. Percorrerá também uma investigação sobre

¹ "Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso." Haykin (2001, p. 32)

aplicações das redes neurais em produtos audiovisuais e examinará possibilidades e potencialidades de uso em obras de arte digital visual e jogos digitais.

II. METODOLOGIA DA PESQUISA

A abordagem adotada para este trabalho é a qualitativa, com fins exploratórios usando os instrumentos bibliográficos sobre o tema, investigando: o aplicativo Google Tradutor na versão *web* e versão para o sistema operacional Android para *smartphones*, o aplicativo Quick Draw! versão *web*, o aplicativo *AutoDraw* versão *web*, o aplicativo Neural Talk 2 versão *web*, o aplicativo Deep Dream versão *web*.

III. REFERENCIAS TEÓRICAS

A. O contexto tecnológico

Com a popularização dos *smartphones*, o grande público tem em mãos uma potente ferramenta em constante evolução. Os softwares e plataformas são atualizadas regularmente, abastecidos com novas tecnologias, oriundas de pesquisas de grandes empresas que buscam despontar como pioneiros nas soluções e ferramentas digitais, transferindo trabalhos estruturados realizados por humanos, sejam repetitivos ou analíticos, para responsabilidade dos mecanismos de inteligência artificial. (SON, 2017)

É possível que as ferramentas de tradução sejam cada vez mais utilizadas, servindo usuários da rede mundial de computadores de diversas partes do mundo em busca de informações específicas que muitas vezes não estão no idioma de compreensão do mesmo.

A ferramenta Google Tradutor, que está integrada às soluções e aplicativos da empresa americana Google Inc. e é utilizada por 500 milhões de pessoas ao redor do mundo,

desponta no Brasil. O país é o que mais utiliza esse serviço. (TUROVSKY, 2016)

B. Redes neurais artificiais

Quando a vontade de aproximar o comportamento dos *softwares* à forma de processamento da mente humana tornou-se motivadora de estudos, foi inferido que uma das formas de alcançar esse feito seria espelhar no meio digital a estrutura comportamental de processamento de informação do cérebro biológico. Surgiram dessas análises as Redes Neurais Artificiais. (PORTUGAL;FERNANDES, 1996)

Na Figura 1.1 é possível observar a estrutura de processamento do neurônio humano, que foi utilizada como base para o desenho das estruturas de Redes Neurais Artificiais. Os estímulos chegam aos detritos que em seguida são transmitidos pelo axônio para as terminações sinápticas e os resultados de saída são passados para outro neurônio e assim sucessivamente. As Redes Neurais Artificiais podem ter uma ou múltiplas camadas. Na versão artificial de rede neural de camada única, as entradas são consideradas variáveis com peso e o produto da soma dos sinais (entradas) é passada para uma função agregadora, que em seguida reage com a uma função de ativação que produzirá como resultado o sinal de saída. (BRANQUINHO, 2013)

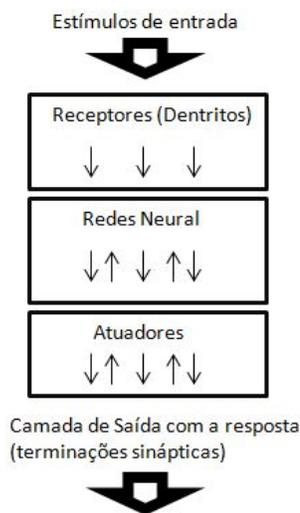


FIGURA 1.1 Representação em diagrama do funcionamento de um neurônio biológico. HAYKIN (2001)

Um exemplo de Rede Neural Artificial de camada única é a Rede Perceptron criada por Frank Rosenblatt em 1958 (ZAMBIASI, 2002, p.10). “O Perceptron foi criado em

1958 por Rosenblatt, sendo a forma mais simples da configuração de uma rede neural artificial, uma vez que é constituída de uma única camada neural e de um único neurônio.” (PALMIERE, 2016) Essas redes de única camada são utilizadas para solucionar “problemas linearmente separáveis”(ZAMBIASI, 2002, p.12).

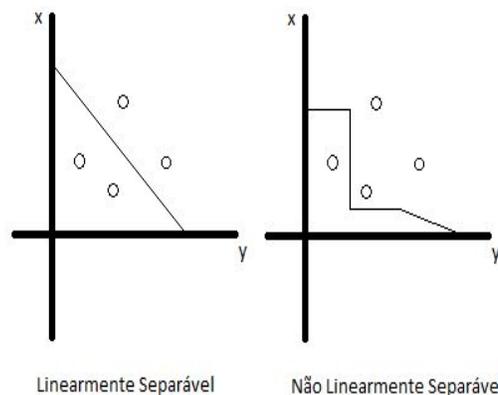


FIGURA 1.2 Representação de problema Linearmente Separável e Não Linearmente Separável. PALMIERE(2016)

Quando a rede neural possui mais de uma camada, ela é denominada rede neural de múltiplas camadas (Figura 1.3). Elas atenderão a necessidade de soluções para problemas não linearmente separáveis. Em cada camada da rede neural haverá vários neurônios. (BRANQUINHO, 2013) Cada camada possui uma função específica e as intermediárias, também denominadas de ocultas fazem o papel de função agregadora, “codificando através de seus pesos as características apresentadas nos padrões de entrada, fazendo com que a rede crie uma representação própria do problema, com mais riqueza e complexidade”. (OLIVEIRA, 2008, p.16)

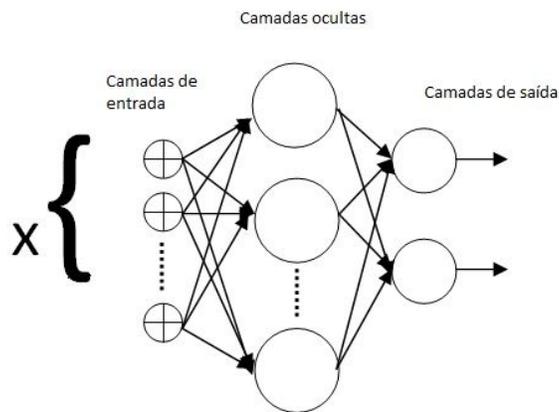


FIGURA 1.3 Representação em diagrama do funcionamento de uma rede neural multicamadas. Os círculos

representam os neurônios e as arestas as sinapses. Goldshimit (2010, p. 83)

C. *Aprendizagem de uma rede neural*

É relevante ponderar que “a propriedade que é de importância primordial para uma rede neural é sua habilidade de aprender a partir de seu ambiente e de melhorar o seu desempenho através da aprendizagem.” (HAYKIN, 2001, p. 75)

As diferenças entre os tipos de algoritmos de aprendizagem de redes neurais artificiais estão na maneira que a regulação de pesos acontece. Colocar em prática um algoritmo de aprendizagem em uma rede neural “consiste em apresentar à rede padrões de informação num processo iterativo onde os pesos sinápticos são ajustados, adquirindo dessa maneira o conhecimento do ambiente no qual irá operar.” (OLIVEIRA, 2008, p.18)

Segundo Goldshimit (2010), o processo de aprendizado possui basicamente os seguintes passos:

- Estímulo do ambiente para receber um padrão de entrada retirado de um conjunto histórico de padrões ou dados;
- Ajustar nos pesos sinápticos;
- A rede neural passa a responder de uma nova forma aos dados e padrões de entrada;
-

D. *Reconhecimento automático de caracteres*

Para realizar a tradução de um termo ou frase, o Google utiliza uma fonte de dados licenciada por idioma² e uma tecnologia de reconhecimento automático de caracteres que anteriormente era utilizada no aplicativo Google Goggles³.

Também denominado de OCR⁴, o reconhecimento automático de caracteres “utiliza métodos mais expressos e convencionais de reconhecimento da forma das letras”. (VALLE JR, 2003, p.45)

Segundo PATEL(2012), o sistema computadorizado de OCR tem sido adotado em diversos tipos de aplicações, como leitura de placas de veículos, reconhecimento de caracteres em imagens reais e extração de texto de documentos impressos. Em todos eles o sistema pode se deparar com problemas. Para um computador diferenciar o número “0” da letra “O”, é um desafio. Outro problema é

extrair caracteres que possuem plano de fundo muito escuro ou estão impressos em gráficos. Para calibrar a assertividade do OCR o sistema recebe uma porcentagem, limite de confiabilidade. Dependendo do retorno em porcentagem o sistema determina se aquele caractere foi reconhecido com sucesso. Hoje a assertividade do OCR está entre 71% e 98%.

Para reconhecer os caracteres em uma imagem, o Google Tradutor captura uma imagem, converte-a em preto e branco, elimina sombras na imagem e passa por filtros que contribuem para que a imagem se limite em exibir formas em preto e branco⁵. Como resultado, existirão agrupamentos de elementos. As manchas escuras que estiverem próximas umas das outras podem ser caracteres. Esses potenciais caracteres são aplicados a um classificador de caracteres que retorna com uma análise de confiabilidade de identificação. Com base nessa análise, é possível descartar informação irrelevante e concentrar a informação com caracteres reconhecidos para serem transmitidos a próxima etapa da tradução.⁶

IV. ANÁLISE DO GOOGLE TRADUTOR

A. *História da ferramenta*

O Google Tradutor é uma ferramenta gratuita, multiplataforma, que está incorporada em outros serviços de sua empresa desenvolvedora, a americana Google, como *browsers* e aplicativos de comunicação, com capacidade de traduzir em até 72 linguagens.⁷ A ferramenta também é oferecida a desenvolvedores e empresas que desejam incorporar a funcionalidade em seus sistemas por meio de uma integração paga.⁸

A ferramenta, que começou a ser desenvolvida em 2001, é utilizada por 200 milhões de pessoas por dia, sendo que 92% dos usuários acessam o serviço fora dos Estados Unidos. Em 2006 quando foi disponibilizada para o público, contava com quatro idiomas para tradução.⁹

Segundo SCHUSTER e LE (2016), em seu primórdio o Google Tradutor utilizava uma tradução automática baseada

⁵ Também chamada de imagem binarizada.

⁶ GOOD, Otávio. Can Your Phone Translate Signs On The Fly?. Youtube, 09 abr. 2017. Disponível em <<https://www.youtube.com/watch?v=0zKU7jDA2nc>>. Acesso em: 09 de abr. 2017.

⁷ About Google Translate. <https://translate.google.com/about/intl/en_ALL/languages/>. Acesso em: 20 de abr. 2017.

⁸ Tools for Business - Google Translate. <http://translate.google.com/about/intl/en_ALL/forbusiness.html>. Acesso em: 20 de abr. 2017.

⁹ Google Translate now serves 200 million people daily. <<https://www.cnet.com/news/google-translate-now-serves-200-million-people-daily/>>. Acesso em: 20 de abr. 2017.

²Google Translate: Open source components and licenses. Disponível em <http://translate.google.com/about/intl/en_ALL/license.html> Acesso em: 09 de abr. 2017.

³ 4 Amazing Uses Of Google Goggles for Android. Disponível em <<http://www.guidingtech.com/11682/amazing-uses-of-google-goggles-for-android/>> Acesso em: 09 de abr. 2017.

⁴ Do inglês: *Optical Character Recognition*.

em frases, denominada *Phrase-Based Machine Translation*¹⁰. Em 2012, ela incorporou a funcionalidade de reconhecimento de voz, para aparelhos que possuem microfone, que permite o reconhecimento da palavra ou frase proferida por um ser humano e sua tradução posterior para o idioma desejado. Essa funcionalidade apresentava erros abaixo de 20% no reconhecimento. Nota-se também que nessa época incorporou-se pesquisas em redes neurais para tornar o sistema mais acurado.¹¹

Em 2014 pesquisas em redes neurais trouxeram para o Google Tradutor, dentre outros produtos da Google, a interpretação de caracteres provenientes de imagens capturadas por câmeras.¹²

B. Google Neural Machine Translation

Foi anunciado o sistema Google Neural Machine Translation (GNMT) em novembro de 2016. Com a adição dessa tecnologia ao Google Tradutor, a equipe de desenvolvedores espera paulatinamente superar as fraquezas do sistema e atingir melhores resultados em precisão e velocidade de tradução. (SCHUSTER; LE,2016) A otimização advinda com a adoção do sistema GNMT consiste na diminuição do tempo de aprendizagem, com a priorização de tradução de palavras mais usuais e prováveis, deixando a tradução de termos e palavras mais raras para uma instância de processamento de uso menos frequente. (YU, 2016, p.1) Também foi adicionado em novembro de 2016 a capacidade do software do aplicativo traduzir sentenças ou frases ao invés de traduzir sempre palavra por palavra em todas as situações, economizando em processamento e aumentando a assertividade, buscando entender o contexto dos termos para executar a tradução¹³. Observa-se na Figura 1.4 a oferta para o usuário de duas opções de pacote de idioma para tradução instantânea. Um denominado “Pequeno” contempla palavras

de uso frequente, com maior probabilidade de uso. Outro denominado “Grande” contempla maior gama de palavras e oferece uma tradução mais assertiva. O que difere entre os pacotes de tradução é o tamanho em *megabytes* para armazenagem interna do dicionário de tradução, *offline*, no aparelho *smartphone* do usuário. O sistema GNMT reduz a quantidade de erros de tradução de 55% a 85%, utilizando amostras de sites de notícias e artigos do Wikipédia, sendo avaliado por tradutores humanos bilíngues. (SCHUSTER; LE,2016)

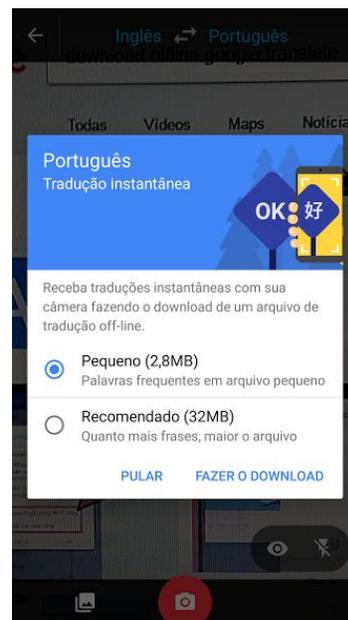


FIGURA 1.4 Tela do aplicativo Google Tradutor exibindo opções de pacotes de idioma para tradução instantânea.

O GNMT possui um tratamento especial para palavras que não constam no vocabulário, como nomes próprios. Essas palavras são simplesmente copiadas para o resultado da tradução final. (YU, 2016, p.7)

O mecanismo de tradução do GNMT ainda recebe intervenções de usuários de todo mundo com o propósito de aprimoramento. Em sua versão incorporada ao mecanismo de busca do Google (Figura 1.5) o Google Tradutor, ele continua aberto a receber sugestões de tradução dos usuários por meio do botão *feedback*. Nessa opção é possível informar ao sistema que a tradução não está correta ou incompleta e sugerir uma tradução fidedigna.

¹⁰ KOEHN, Philipp; OCH, Franz Josef; MARCU, Daniel. Statistical Phrase-Based Translation.

<<http://homepages.inf.ed.ac.uk/pkoehn/publications/phrase2003.pdf>>. Acesso em: 10 de jul. 2017.

¹¹ Speech Recognition and Deep Learning. 2012, Google Research Blog.<<https://research.googleblog.com/2012/08/speech-recognition-and-deep-learning.html>> Acesso em: 20 de abr. 2017.

¹² Building a deeper understanding of images. 2014, Google Research Blog.<<https://research.googleblog.com/2014/09/building-deeper-understanding-of-images.html>> Acesso em: 20 de abr. 2017.

¹³ Found in translation: More accurate, fluent sentences in Google Translate. 2016, Google Blog.<<https://blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/>> Acesso em: 09 de maio 2017.



FIGURA 1.5 Tela do Google Tradutor incorporado no mecanismo de busca do Google.

O Google Tradutor possui também a Comunidade do Google Tradutor, uma plataforma colaborativa, que requer cadastro, onde usuários da ferramenta de todo o mundo são capazes de colaborar com o Google Tradutor traduzindo frases, validando frases realizadas por outros usuários. Ao mesmo tempo, lidando com os dados inseridos pelos usuários, a ferramenta aprimora seu mecanismo de Inteligência Artificial, abastecendo-se do conteúdo gerado por usuários colaboradores.

A partir do clique na opção de escrita à mão, o usuário pode imputar símbolos, ideogramas, letras do alfabeto romano e distintos, originários de outros sistemas de escrita, que imediatamente serão reconhecidos pelo Google Tradutor.¹⁴ O processamento interno sugerirá termos digitalizados do conteúdo imputado à mão, para que, logo em seguida, o aplicativo exponha uma tradução.

C. Tradução instantânea de imagens e áudio

A partir da versão publicada no início de 2015 o aplicativo Google Tradutor incorporou duas funcionalidades de tradução instantânea¹⁵: a tradução utilizando a leitura da câmera do *smartphone* em tempo real (denominada *Word Lens*) e a tradução simultânea de áudio utilizando o microfone do *smartphone*.

O uso do *Word Lens*, lançado em 2010, tem aplicabilidade vasta, embora sua origem tenha sido a necessidade de traduzir placas e impressos em língua estrangeira segundo Otávio Good, um dos desenvolvedores responsáveis pela ferramenta. Com o aplicativo ligado e a opção *Word Lens* escolhida, o usuário tem a possibilidade de mirar a câmera de seu dispositivo para qualquer impresso em

¹⁴ Google Tradutor com entrada manuscrita ou pelo teclado virtual. Disponível em: <<https://support.google.com/translate/answer/6142469?co=GENIE.Platform%3DDesktop&hl=pt-BR>> Acesso em: 30 de maio 2017.

¹⁵ Google libera app de tradução instantânea e de textos em imagens. Disponível em: <<http://link.estadao.com.br/noticias/geral,google-libera-app-de-traducao-instantanea-e-de-textos-em-imagens,10000029822>> Acesso em: 22 de maio 2017.

caracteres não manuscritos, para promover uma tradução instantânea de acordo com idiomas parametrizados, visualizando em tempo real na tela do aparelho a tradução imputada sobre a superfície onde os caracteres foram impressos.¹⁶ O *Word Lens* é uma solução de tradução usando realidade aumentada, tecnologia que sobrepõe imagens digitais em imagens capturadas por câmeras em tempo real,¹⁷ que foi incorporada ao Google Tradutor em 2014. Anteriormente era um aplicativo independente da empresa Quest Visual.¹⁸

V. PRÁTICAS DE USO DE REDES NEURAIS ARTIFICIAIS EM ESTRUTURAS AUDIOVISUAIS

Com o crescente interesse e estudos em redes neurais artificiais, diversos projetos passaram a incorporar mecanismos de inteligência artificial em seus sistemas. Grande parte desses projetos usam da interação com a realidade pelo meio visual, utilizando reconhecimento de imagens e realidade aumentada para interpretar e reagir aos dados capturados. Os trabalhos investigados neste artigo foram escolhidos pela utilização das redes neurais artificiais em suas aplicações e pela disponibilização do uso ou visualização na *internet*.

A interpretação do computador para desenhos produzidos por humanos e a capacidade do computador produzir desenhos como os humanos está presente no projeto envolvendo o jogo *Quick, Draw!*, produzido pelo Google. Com o propósito de criar um banco de dados de desenhos simplificados e promover o aprendizado de sua rede neural, o *Quick Draw!*, jogo *online* desenvolvido pelo *Google Creative Lab* em parceria com pesquisadores de aprendizado de máquina¹⁹, promove um desafio ao jogador: a partir de uma sugestão, fazer um desenho utilizando um pincel digital, durante tempo cronometrado de vinte segundos, com o objetivo da inteligência artificial do jogo decifre o desenho. O jogo já

¹⁶Otávio Good explains Word Len. Disponível em: <<https://www.youtube.com/watch?v=2-t80mHtM-Y>> Acesso em: 22 de maio 2017.

¹⁷About Augmented Reality. Disponível em: <<http://www.augmented.org/blog/about/>> Acesso em: 18 de junho 2017.

¹⁸ Google acquires Quest Visual to integrate its Word Lens augmented reality technology into Google Translate. Disponível em: <https://thenextweb.com/google/2014/05/16/google-acquires-quest-visual-incorporate-word-lens-augmented-reality-technology-google-translate/#.tnw_qm9xoNPE> Acesso em: 22 de maio 2017.

¹⁹ See 111,000 attempts to draw the Mona Lisa in under 20 seconds, collected by Google. Disponível em: <<https://qz.com/987159/google-googl-quick-draw-50-million-images-of-peoples-doodles/>> Acesso em: 28 de maio 2017.

conta com uma biblioteca de mais de 50 milhões de desenhos salvos que, pelo aprendizado, interferem na forma que o jogo irá interpretar os próximos desenhos inseridos no jogo.²⁰ Em sua base de dados²¹, o *Quick, Draw!* reúne milhares de desenhos realizados pelos jogadores, separados pelos termos sugeridos pelo jogo para a realização dos desenhos. A inteligência artificial presente no jogo foi denominada pelos desenvolvedores como Sketch-RNN, uma rede neural treinada por seres humanos para reconhecimento de desenhos de objetos comuns. (HA; ECK, 2017, p.1) O código da rede neural é *open source* e está disponível *online* para utilização de desenvolvedores interessados²². Os frutos dessa pesquisa proporcionam a capacidade do Sketch-RNN em aprimorar a construção de desenhos vetoriais originais. Na Figura 1.6, é possível observar a construção da rede neural de desenhos vetoriais a partir de esboços humanos identificados como representações de gatos.²³

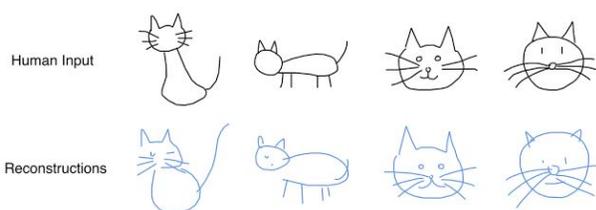


FIGURA 1.6²⁴ Reconstruções do Sketch-RNN para desenhos humanos representando gatos.

O mecanismo de reconhecimento de desenhos do Sketch-RNN é o mesmo que é utilizado para reconhecimento de ideogramas e caracteres desenhados no aplicativo Google Tradutor.²⁵

O sistema *DeepDream* é um programa de processamento de imagens criado pelo Google, que realiza, por meio de redes neurais artificiais, releituras de fotos após a interpretação por

meio do computador, dos objetos e seres encontrados na imagem.²⁶ Segundo pesquisadores do Google, o programa apresenta representações visuais do processamento das redes neurais, sendo notável o aprendizado e a classificação que os algoritmos realizam, além de proporcionar imagens com estética artística, próximo do abstrato.²⁷ A ferramenta pode ser utilizada por visitantes do *site* após cadastro. Existem dois tipos de processamento, o *Deep Style*, que proporciona um efeito pré-definido sobre a imagem (resultado observado na Figura 1.7) e o *Deep Dream*, que exerce alteração sem estilo, proporcionada pelos algoritmos das redes neurais do programa. Para realizar o processamento, basta ao usuário fazer o *upload* de uma foto de sua escolha. O usuário pode escolher manter a imagem privada para apenas seu acesso pessoal ou permitir que a foto processada faça parte da galeria pública do programa²⁸. A imagem gerada é exibida logo em seguida ao processamento.



FIGURA 1.7 Imagem resultante do processamento utilizando o *Deep Style* do programa *DeepDream*.

O código fonte do projeto *DeepDream*, para execução e colaboração de programadores, está disponível na *internet*.²⁹ A ferramenta conta com uma conta pública na rede social de fotos *Instagram* com uma seleção de imagens processadas pelo programa.³⁰

²⁰ Quick Draw. <<https://quickdraw.withgoogle.com/>> Acesso em: 28 de maio 2017.

²¹ Quick Draw Data Set. Disponível em: <<https://quickdraw.withgoogle.com/data/>> Acesso em: 28 de maio 2017.

²² Sketch-RNN model released in Magenta. Disponível em: <https://magenta.tensorflow.org/sketch_rnn> Acesso em: 28 de maio 2017.

²³ Teaching Machines to Draw, Google Research Blog. Disponível em: <<https://research.googleblog.com/2017/04/teaching-machines-to-draw.html>> Acesso em: 28 de maio 2017.

²⁴ Fonte da Figura: Teaching Machines to Draw, Google Research Blog. Disponível em: <<https://research.googleblog.com/2017/04/teaching-machines-to-draw.html>> Acesso em: 28 de maio 2017.

²⁵ A.I. Experiments: Quick, Draw! Disponível em: <<https://www.youtube.com/watch?v=X8v1GWzZYJ4>> Acesso em: 28 de maio 2017.

²⁶ Google Deep Dream: tecnologia gera imagens com resultados psicodélicos. Disponível em: <<http://www.techtudo.com.br/noticias/noticia/2015/07/google-deep-dream-tecnologia-gera-imagens-com-resultados-psicodelicos.html>> Acesso em: 30 de maio 2017.

²⁷ DeepDream - a code example for visualizing Neural Networks, Google Research Blog. Disponível em: <<https://research.googleblog.com/2015/07/deepdream-code-example-for-visualizing.html>> Acesso em: 30 de maio 2017.

²⁸ DeepDream. Disponível em: <<https://deepdreamgenerator.com>> Acesso em: 30 de maio 2017.

²⁹ Google/DeepDream. Disponível em: <<https://github.com/google/deepdream>> Acesso em: 30 de maio 2017.

³⁰ DeepDream Instagram. Disponível em:

O *NeuralTalk 2* é uma sistema de reconhecimento e identificação de imagens estáticas e provenientes de capturas em vídeo, elaborado por uma equipe de desenvolvedores da Universidade de Stanford, na linguagem de programação Lua, usando o *framework* Torch e com código livre sendo executado em sistemas *Ubuntu*.³¹ O *NeuralTalk 2* é a evolução do sistema *NeuralTalk* que foi descontinuado em novembro de 2015.³² A proposta do *NeuralTalk 2* é prover descrição automática de conteúdo de imagem correspondendo dados visuais aos termos linguísticos no idioma inglês. Para treinamento de suas redes neurais de identificação de imagens, foram utilizados os bancos de dados *onlines* de imagens Flickr8K, Flickr30K e o MSCOCO da Microsoft. O processamento do *NeuralTalk 2* permite a descrição de itens que compõem a imagem e a sua concatenação, permitindo a formação de frases com o conteúdo exibido na imagem. O sistema de qualificação do resultado obtido pelas redes neurais é pontuado em BLEU (PAPINENI et al., 2002), método de avaliação de traduções realizador por máquina, sendo que, quanto mais positivo é o valor obtido em pontuação BLEU, maior é a confiabilidade do sistema no resultado obtido (KARPATHY; FEI-FEI, 2017)

Em uma experiência realizada com o processamento de um vídeo capturado por câmera, é possível visualizar os resultados gerados pelas redes neurais, expondo a descrição do que a inteligência artificial percebe das cenas gravadas.³³

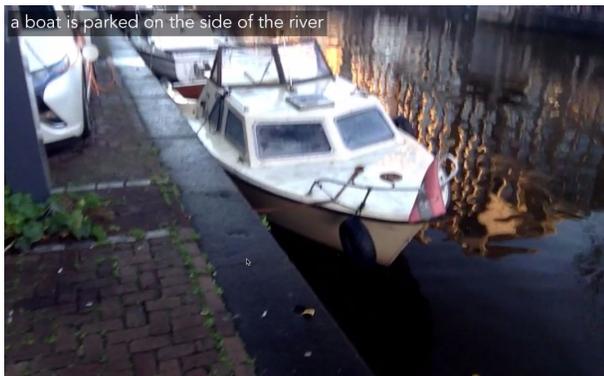


FIGURA 1.8 *Frame* do vídeo “*NeuralTalk and Walk*” com a legenda gerada pelo *NeuralTalk 2* “*a boat is parked on the side of the river*”.

Durante a exibição de legendas, é possível isolar capturas

<<https://www.instagram.com/deepdreamgenerator/>> Acesso em: 30 de maio 2017.

³¹ *NeuralTalk2* < <https://github.com/karpathy/neuraltalk2>> Acesso em: 06 de junho 2017.

³²*NeuralTalk*. Disponível

em:<<https://github.com/karpathy/neuraltalk>> Acesso em: 06 de junho 2017.

³³ *NeuralTalk and Walk*. Disponível

em:<<https://vimeo.com/146492001>> Acesso em: 06 de junho 2017.

exatas como a da Figura 1.8, onde se lê “um barco estacionado ao lado do rio”.³⁴ Também é possível isolar capturas onde a descrição do conteúdo da imagem foi equivocada, como no caso da Figura 1.9, que descreve “ um foco em um telefone celular sobre a mesa”³⁵, onde na verdade é visualizado a traseira de um veículo automotor e um aglomerado de bicicletas.



FIGURA 1.9 *Frame* do vídeo “*NeuralTalk and Walk*” com a legenda gerada pelo *NeuralTalk 2* “*a close up of a cell phone on a table*”.

Em uma plataforma online de exposição de códigos, a *DeepIA*, usuários de todo o mundo podem executar o *NeuralTalk 2* realizando *uploads* de imagens existentes nos computadores pessoais.³⁶ Após o *upload* de imagem única e o processamento das redes neurais, uma descrição em texto é exibida como resultado. Na página *NeuralTalk 2 Web Demo*³⁷, uma galeria de imagens processadas provenientes do banco de imagens *online* MSCOCO são exibidas com as descrições geradas pelas redes neurais do sistema.

Em apresentação realizada no evento *Deep Learning School*³⁸, um dos desenvolvedores do sistema, Andrej Karpathy, citou como aplicações para o reconhecimento e geração automática de descrição de imagens: aprimoramento de carros autônomos; reconhecimento facial; reconhecimento de números de logradouros; classificação e organização dos resultados do Google Imagens; diagnósticos de exames

³⁴ “*a boat is parked on the side of the river*”, Tradução do autor.

³⁵ “*a close up of a cell phone on a table*”, Tradução do autor.

³⁶*DeepAI, NeuralTalk 2*. Disponível em:<<https://deepai.org/machine-learning-model/neuraltalk>> Acesso em: 06 de junho 2017.

³⁷*NeuralTalk 2 Web Demo* . Disponível em:<<http://cs.stanford.edu/people/karpathy/deepimagesent/generationdemo/>> Acesso em: 06 de junho 2017.

³⁸ *Deep Learning School*. Disponível em: <https://www.bayareadschool.org/> Acesso em: 06 de junho 2017.

médicos; reconhecimento de caracteres; reconhecimento de baleias capturadas por fotos do oceano; análises de imagens capturadas por satélite; exploração de imagens de galáxias; criação de legendas; aumento na *performance* de inteligência artificial aplicada em jogos eletrônicos; desenvolvimento de aplicações artísticas.³⁹

Disponível na *internet*, o *AutoDraw*⁴⁰ é uma aplicação para produção de desenhos vetoriais. Utilizando inteligência artificial para o reconhecimento de formas desenhadas usando o cursor do aparelho (computador, *smarthphone*, etc) em uma tela, a aplicação sugere desenhos vetoriais prontos, da própria base de dados, deixando a cargo do usuário a escolha e formatação. O *AutoDraw* é uma produção dos desenvolvedores Dan Motzenbecker e Kyle Phillips e colaboradores do Google Creative Lab. Com a proposta de facilitar a produção de artes visuais, o *AutoDraw* contempla uma ferramenta pincel que expõe vetores semelhantes aos desenhos realizados pelo usuário, ferramenta de inclusão de texto, formas geométricas, diversidade de cores, e um pincel para desenho livre, sem associação com formas vetorizadas. O usuário pode redimensionar o tamanho da tela e escolher entre o *layout* horizontal ou vertical. Após a conclusão dos desenhos, o usuário tem a opção de realizar o *download* da imagem no formato PNG para seu aparelho digital. A inteligência artificial utilizada para o reconhecimento dos desenhos e a sugestão de vetores é a mesma utilizada na aplicação *Quick Draw!*. É possível, para usuários interessados, submeter desenhos vetorizados para contribuir com a base de dados da aplicação. Para a submissão, é necessário preencher um formulário e adequar o desenho a um padrão de formato pré-estabelecido pelos autores.⁴¹

VI. CONSIDERAÇÕES FINAIS

A aplicação de redes neurais artificiais em aplicativos de interação humana ampliam a capacidade de absorção de informação e produtividade por parte dos usuários. A inteligência artificial sobrepasa o conceito de facilitar as atividades humanas e elucida para os pesquisadores alguns processos da mente humana. Com a ferramenta de câmera do Google Tradutor, o usuário experimenta uma realidade alternativa que diminui lacunas de comunicação entre povos com línguas diferentes. Sua interface simplificada combinada a fácil utilização a imediata resposta torna a ferramenta prática

para ser efetivamente usada e experimentada. A realidade aumentada é aplicada com precisão, tornando o processo de tradução amigável e prazeroso. Diversas informações passam a ser consumidas por usuários não dotados do domínio da leitura do idioma original da informação sem a interferência de um tradutor humano ou de uma árdua tarefa de tradução utilizando um dicionário. Caso a tecnologia de óculos inteligentes entre em voga novamente, esse é um serviço que pode ser incorporado ao *gadget*.

A utilização da mesma inteligência artificial do aplicativo *Quick Draw!* no aplicativo *AutoDraw* é um exemplo da convergência que pode ser realizada entre as aplicações, deixando claro que uma rede neural treinada em um ambiente pode ser aplicada em outro, com objetivos diferentes.

A aplicação *AutoDraw* oferece um serviço com intuito de ajudar pessoas a produzirem peças visuais digitais em tempo reduzido e de forma prática. Esse serviço abre possibilidades de produção de peças visuais bem acabadas mas reduzidas em originalidade. Pode ser interpretado que ela reduz a expressão criativa na concepção de uma obra visual. A imperfeição a inexatidão acompanha toda a história da produção artística humana.

A dedicação das redes neurais para interpretação de imagens capturadas por câmera desperta inúmeras possibilidades no campo de automação. Existe aplicação para essa tecnologia em carros autônomos, sistemas de segurança e identificação de pessoas, serviços de acessibilidade, robótica, entre outras. É possível vislumbrar com o *NeuralTalk 2* o entendimento a partir do computador de itens isolados e dentro de contextos.

Toda interação humana é importante para o constante aprimoramento das redes neurais com o aprendizado promovido pelos algoritmos. O *feedback* é crucial para a velocidade de aprendizado e o entendimento do raciocínio humano para a inteligência artificial. As redes neurais artificiais trazem progresso aos resultados do processamento computacional e a acurácia tende a se aprimorar em uma velocidade intensa nos próximos anos.

Notavelmente neste trabalho abordaram-se as redes neurais contribuindo em sistemas que geram resultados desassociados com as identidades dos indivíduos. As redes neurais também podem ser aplicadas para entender comportamentos individuais e oferecer respostas exclusivas para os interatores.

Em meio a tudo isso, os artistas digitais encontram ferramentas novas para experimentar e linguagens para conhecer. A aplicação *DeepDream* é um bom exemplo com extensas ferramentas para representação, releitura de imagens e construção de novos processos criativos. Não somente na produção de obras utilizando as redes neurais no processamento das imagens, outro campo onde os artistas podem tecer arte é a da interferência nos processos das aplicações e deturpação das redes neurais e crítica aos resultados obtidos. Surge junto com as produções de obras artísticas digitais indagações sobre o papel do computador e da inteligência artificial com autores da arte digital. Todos esses

³⁹ *Deep Learning for Computer Vision* (Andrej Karpathy, *OpenAI*), *Deep Learning School* 24 e 25 de Setembro de 2016 . Disponível

em: <<https://www.youtube.com/watch?v=u6aEYuemt0M>>
Acesso em: 06 de junho 2017.

⁴⁰ *AutoDraw*. Disponível em: <<https://www.autodraw.com/>>
Acesso em: 17 de junho 2017.

⁴¹ *Fast Drawing for Everyone*. Disponível em:
<<https://blog.google/topics/machine-learning/fast-drawing-everyone/>> Acesso em: 17 de junho 2017.

campos são extensos e estão livres para serem explorados por pesquisadores.

- [1] BARBOSA, A. H.; FREITAS, M. S. R; NEVES, F. A. Confiabilidade estrutural utilizando o método de Monte Carlo e redes neurais. Rem: Rev. Esc. Minas [online]. 2005, vol.58, n.3, pp.247-255.
- [2] BRANQUINHO, A. Redes Neurais Artificiais (Introdução) – Parte 01, 2013. Disponível em: <<https://www.youtube.com/watch?v=vbf4IzvXvuM&t=1s>>. Acesso em 13 de março de 2017.
- [3] BRANQUINHO, A. Redes Neurais Artificiais (Adaline, Algoritmos e Considerações Finais) - Parte 03, 2013. Disponível em: <<https://www.youtube.com/watch?v=XtefCjRIWdo&t=39s>>. Acesso em 27 de março de 2017.
- [4] GOLDSHIMITH, Ronaldo R. Uma Introdução à Inteligência Computacional: fundamentos, ferramentas e aplicações. 2010.
- [5] HA, David; ECK Douglas. A Neural Representation of Sketch Drawings, 2017, Cornell Univesity Library. Disponível em: <<https://arxiv.org/pdf/1704.03477.pdf>> Acesso em: 28 de mai. 2017.
- [6] KARPATY, Andrej; FEI-FEI, Li. Deep Visual-Semantic Alignments for Generating Image Descriptions. Department of Computer Science, Stanford University, 2017.
- [7] OLIVEIRA, Angelo Rocha. Redes Neurais Artificiais Aplicadas na Detecção, Classificação e Localização de Defeitos em Linhas de Transmissão, 2005.
- [8] PALMIERE, Sergio E.. Rede Perceptron de uma única camada. Disponível em: <<https://www.embarcados.com.br/rede-perceptron-de-uma-unica-camada/>> Acesso em: 12 de mar. 2017.
- [9] PAPINENI, K.; ROUKOS, S.; WARD, T.; ZHU, W. (2002). BLEU: a method for automatic evaluation of machine translation. In Proceedings of ACL-2002. Philadelphia, PA. pp. 311-318
- [10] PATEL, C., PATEL, A. and PATEL, D. (2012) Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study. International Journal of Computer Applications, 55, 50-56.
- [11] PORTUGAL, M. S.; FERNANDES, L. G. L. Redes neurais artificiais e previsão de séries econômicas: uma introdução. Nova Economia, v. 6, n. 1, 1996.
- [12] SIMON, Haykin. Redes Neurais:Princípios e Prática. Bookman, 2001.
- [13] SCHUSTER, M.; LE, V. Quoc. A Neural Network for Machine Translation, at Production Scale. 2016, Google Research Blog. Disponível em: <<https://research.googleblog.com/2016/09/a-neural-network-for-machin-e.html>>. Acesso em 20 de abril de 2017.
- [14] SON, Hugh. JPMorgan Software Does in Seconds What Took Lawyers 360,000 Hours, 2017. Disponível em: <<https://www.bloomberg.com/news/articles/2017-02-28/jpmorgan-mars-hals-an-army-of-developers-to-automate-high-finance>> Acesso em: 12 de mar. 2017.
- [15] TUROVSKY, Barak. Ten years of Google Translate, 2016. Disponível em: <<https://blog.google/products/translate/ten-years-of-google-translate>> Acesso em: 12 de mar. 2017.
- [16] VALLE JR., E. A., Sistemas de informação multimídia na preservação de acervos permanentes. 2003, Dissertação (Mestrado em Ciência da Computação) – Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Belo Horizonte.
- [17] WU, Y. et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine. 2016, Cornell Univesity Library. <<https://arxiv.org/abs/1609.08144>>Acesso em: 09 de maio 2017.
- [18]
- [19] ZAMBIASI,Saulo P. Introdução as Redes Neurais Artificiais., Disponível em: <<http://www.gsigma.ufsc.br/~popov/aulas/ia/modulo9.pdf>>. Acesso em 20 de março de 2017.
- [20]