



# Deep Learning for Accessibility

Detection and Segmentation of Regions of Interest for  
Sign Language Recognition Systems

---

Rúbia Guerra

Orientador: Prof. Dr. Frederico Gadelha

Coorientadora: Ma. Tamires Rezende

24 de Junho de 2019

Escola de Engenharia, UFMG

1. Introdução
2. Fundamentos
3. Aplicações
4. Deep Learning na Sociedade
5. Conclusão

# Introdução

---

**466 milhões** de pessoas ao redor do mundo apresentam perda auditiva debilitante [WHO, 2019]

Kuenburg et al. [2016]: **comunicação** é uma das maiores barreiras para acesso a serviços de saúde pela comunidade surda

## NOME DA GÍRIA: 007

Significado: Pessoa esperta, inteligente, sedutora, malandra ou com boa lábia

MUNDO  
estranho



**1** Mão direita em "o", com palma voltada para a esquerda



**2** Deslizar a mão para a direita, mantendo a mesma posição em "o"



**3** Com a mão na altura do ombro, fazer o sinal de "7"

Fonte: *Hirata [2019]*

Tecnologias vêm sendo estudadas para auxiliar na comunicação entre pessoas surdas e não-surdas:

## Português-Libras

HandTalk [2019], VLibras [2019].

## Libras-Português

Problema em aberto: Almeida et al. [2014], Rezende [2016], Filho et al. [2017], Amaral et al. [2019]

Er-Rady et al. [2017]:

- Complexidade visual das línguas de sinais

Er-Rady et al. [2017]:

- Complexidade visual das línguas de sinais
- Poucas bases de dados documentadas



Er-Rady et al. [2017]:

- Complexidade visual das línguas de sinais
- Poucas bases de dados documentadas
- Teoria linguística pouco desenvolvida

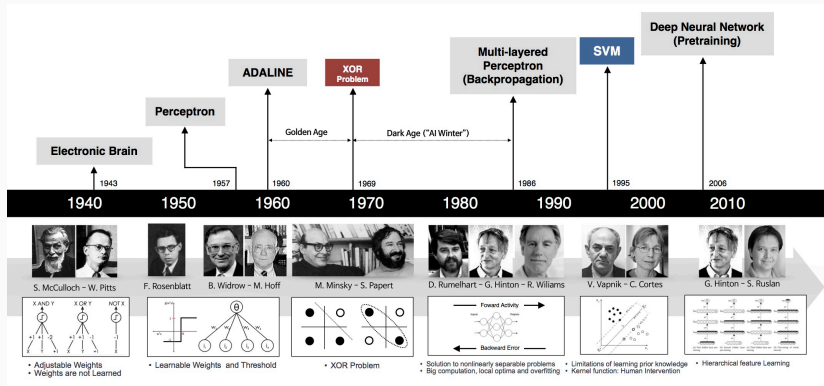
Er-Rady et al. [2017]:

- Complexidade visual das línguas de sinais
- Poucas bases de dados documentadas
- Teoria linguística pouco desenvolvida

# Fundamentos

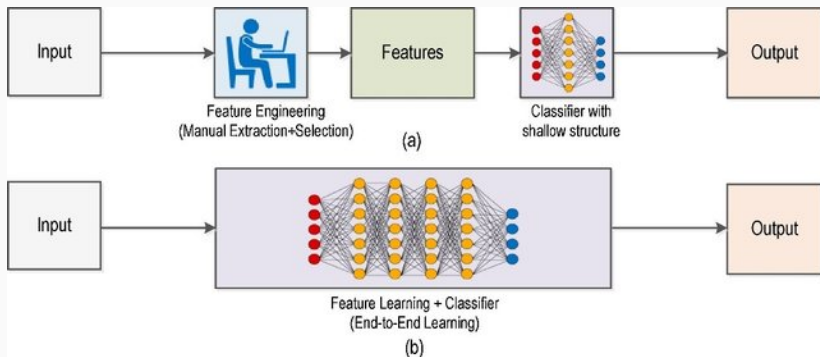
---

# Resumo Histórico



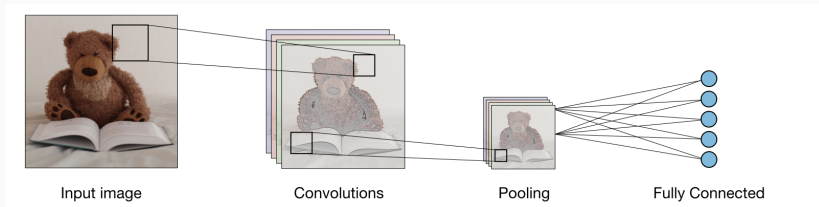
Fonte: *Beam [2017]*

# Deep Learning



Fonte: Wang et al. [2018]

# Redes Neurais Convolucionais

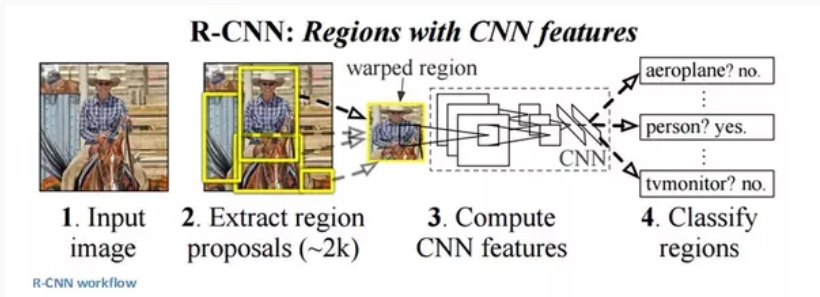


Fonte: *Amidi and Amidi [2018]*

# Aplicações

---

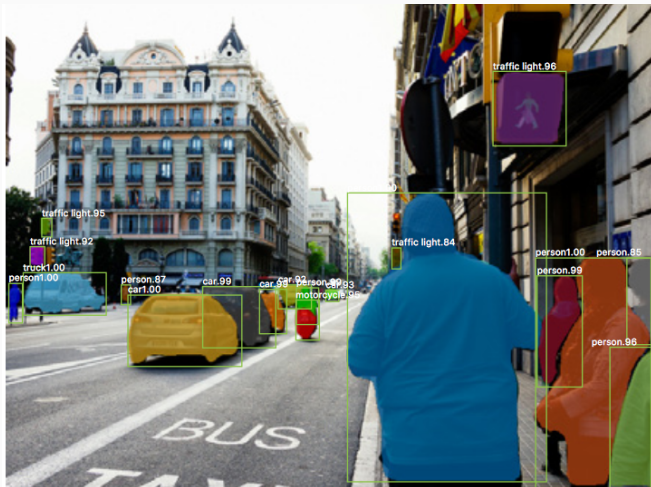
# Segmentação Semântica: R-CNN



Fonte: Girshick et al. [2013]

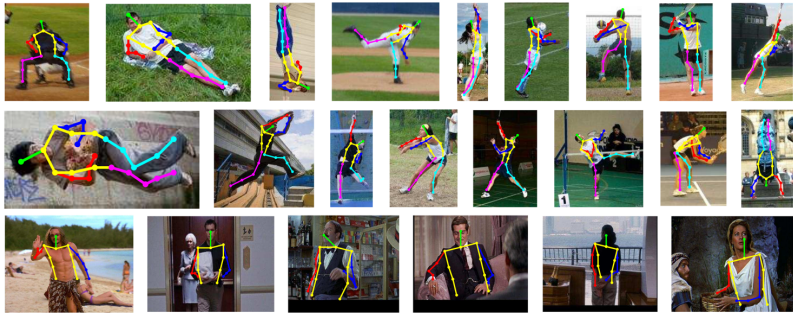


# Segmentação Semântica: Mask R-CNN



Fonte: *Farooq [2018]*

# Estimação de Pose



Fonte: Babu [2019]

# Deep Learning na Sociedade

---



## APPLICATIONS OF DEEP LEARNING

Natural Language  
Processing

Drug Discovery & better  
diagnostics of diseases  
in Healthcare

Robots and Self-  
Driving Cars



Image  
Recognition



Portfolio Management  
& prediction of stock  
price movements.



Speech  
Recognition



Fonte: *Bisht [2019]*



Fonte: *Huq [2019]*

Modelos de Deep Learning utilizam grandes quantidades de dados, por vezes contendo informações sensíveis. Sistemas de IA devem prezar pela segurança do modelo e privacidade dos usuários [Bae et al., 2018]:

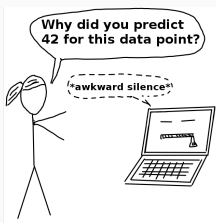
## Secure AI

Proteção contra ataques de *poisoning* e evasão.

## Private AI

Preservação da privacidade dos dados.

Modelos **caixa-preta** são problemáticos em aplicações sensíveis.  
**Interpretabilidade** pode ser atingida através de [Lipton, 2016]:



Fonte: Molnar [2019]

## Transparência

Do modelo, de componentes individuais ou do algoritmo.

## Post-hoc

Visualizações das representações aprendidas, explicações em linguagem natural.

# Conclusão

---

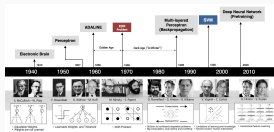


## Motivação

466 milhões de pessoas ao redor do mundo apresentam perda auditiva debilitante [WHO, 2019]

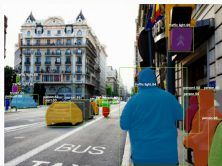
Kuenburg et al. [2016]: **comunicação** é uma das maiores barreiras para acesso a serviços de saúde pela comunidade surda

## Resumo Histórico



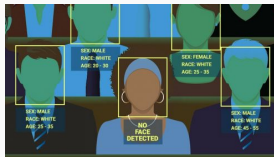
Fonte: Beam [2017]

## Segmentação Semântica: Mask R-CNN



Fonte: Farooq [2018]

## Fairness



Fonte: Huq [2019]

Grata pela presença de todos!

Perguntas?

- S. G. M. Almeida, F. G. Guimarães, and J. A. Ramírez. Feature extraction in brazilian sign language recognition based on phonological structure and using rgb-d sensors. *Expert Systems with Applications*, 41(16):7259–7271, 2014.
- L. Amaral, G. L. N. Júnior, T. Vieira, and T. Vieira. Evaluating deep models for dynamic brazilian sign language recognition. In R. Vera-Rodriguez, J. Fierrez, and A. Morales, editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, page 930–937. Springer International Publishing, 2019. ISBN 978-3-030-13469-3.
- A. Amidi and S. Amidi. Cs 230 - convolutional neural networks cheatsheet. Lecture Notes, Nov 2018. URL <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks>.

- S. C. Babu. A 2019 guide to human pose estimation with deep learning 2019, Apr 2019. URL <https://blog.nanonets.com/human-pose-estimation-2d-guide/>.
- H. Bae, J. Jang, D. Jung, H. Jang, H. Ha, and S. Yoon. Security and privacy issues in deep learning. *arXiv preprint arXiv:1807.11655*, 2018.
- A. L. Beam. Deep learning 101 - part 1: History and background. Online, 2017. URL [https://beamandrew.github.io/deeplearning/2017/02/23/deep\\_learning\\_101\\_part1.html](https://beamandrew.github.io/deeplearning/2017/02/23/deep_learning_101_part1.html). [Accessed June 10th, 2019].
- P. Bisht. Here are some amazing deep learning applications 2019, 2019. URL <https://www.houseofbots.com/news-detail/11881-1-here-are-some-amazing-deep-learning-applications>

- A. Er-Rady, R. Faizi, R. O. H. Thami, and H. Housni. Automatic sign language recognition: A survey. In *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, page 1–7. IEEE, May 2017. ISBN 978-1-5386-0551-6. URL <http://ieeexplore.ieee.org/document/8075561/>.
- U. Farooq. From r-cnn to mask r-cnn, Feb 2018. URL [https://medium.com/@umerfarooq\\_26378/from-r-cnn-to-mask-r-cnn-d6367b196cfd](https://medium.com/@umerfarooq_26378/from-r-cnn-to-mask-r-cnn-d6367b196cfd).
- C. F. F. C. Filho, R. S. de Souza, J. R. dos Santos, B. L. dos Santos, and M. G. F. Costa. A fully automatic method for recognizing hand configurations of brazilian sign language. *Research on Biomedical Engineering*, 33(1):78–89, mar 2017. URL <https://doi.org/10.1590/2446-4740.03816>.

## Referências iv

- R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *arXiv:1311.2524 [cs]*, Nov 2013. URL <http://arxiv.org/abs/1311.2524>. arXiv: 1311.2524.
- HandTalk. Handtalk. Online, 2019. URL <http://www.handtalk.me/>. [Accessed May 31st, 2019].
- G. Hirata. Existem gírias em libras? | superinteressante, 2019. URL <https://super.abril.com.br/cultura/existem-girias-na-lingua-de-sinais-dos-surdos/>.
- P. Huq. Practice makes perfect? ai bias and ml fairness | cctp-607: “big ideas”: Ai to the cloud, 2019. URL <https://blogs.commonsgorgetown.edu/cctp-607-spring2019/2019/03/20/practice-makes-perfect-ai-bias-and-ml-fairness/>.

## Referências v

- A. Kuenburg, P. Fellingner, and J. Fellingner. Health care access among deaf people. *The Journal of Deaf Studies and Deaf Education*, 21(1): 1–10, Jan 2016. ISSN 1081-4159. doi: 10.1093/deafed/env042.
- Z. C. Lipton. The mythos of model interpretability. *arXiv preprint arXiv:1606.03490*, 2016.
- C. Molnar. *Interpretable Machine Learning*. 2019. URL <https://christophm.github.io/interpretable-ml-book/terminology.html>.
- T. M. Rezende. Aplicação de técnicas de inteligência computacional para análise da expressão facial em reconhecimento de sinais de libras. Master's thesis, UFMG, 2016. URL <https://www.ppgee.ufmg.br/defesas/1393M.PDF>.
- VLibras. Vlibras - tradução de português pra libras. Online, 2019. URL <http://www.vlibras.gov.br/>. [Accessed May 31st, 2019].

J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu. Deep learning for smart manufacturing: Methods and applications. *Journal of Manufacturing Systems*, 48:144–156, Jul 2018. ISSN 0278-6125.

WHO. Who | estimates. Online, 2019. URL <http://www.who.int/deafness/estimates/en/>.