Configuration Language for Convolutional Neural Networks

Jony Van Puymbroeck

Promotor: Prof. Dr. Vangheluwe H.

Supervisor: Claudio Gomes

Universiteit Antwerpen

HistoGeneX

Content

- Introduction
- Tool Details
- Tools Used
- Demo
- Related Work

Introduction

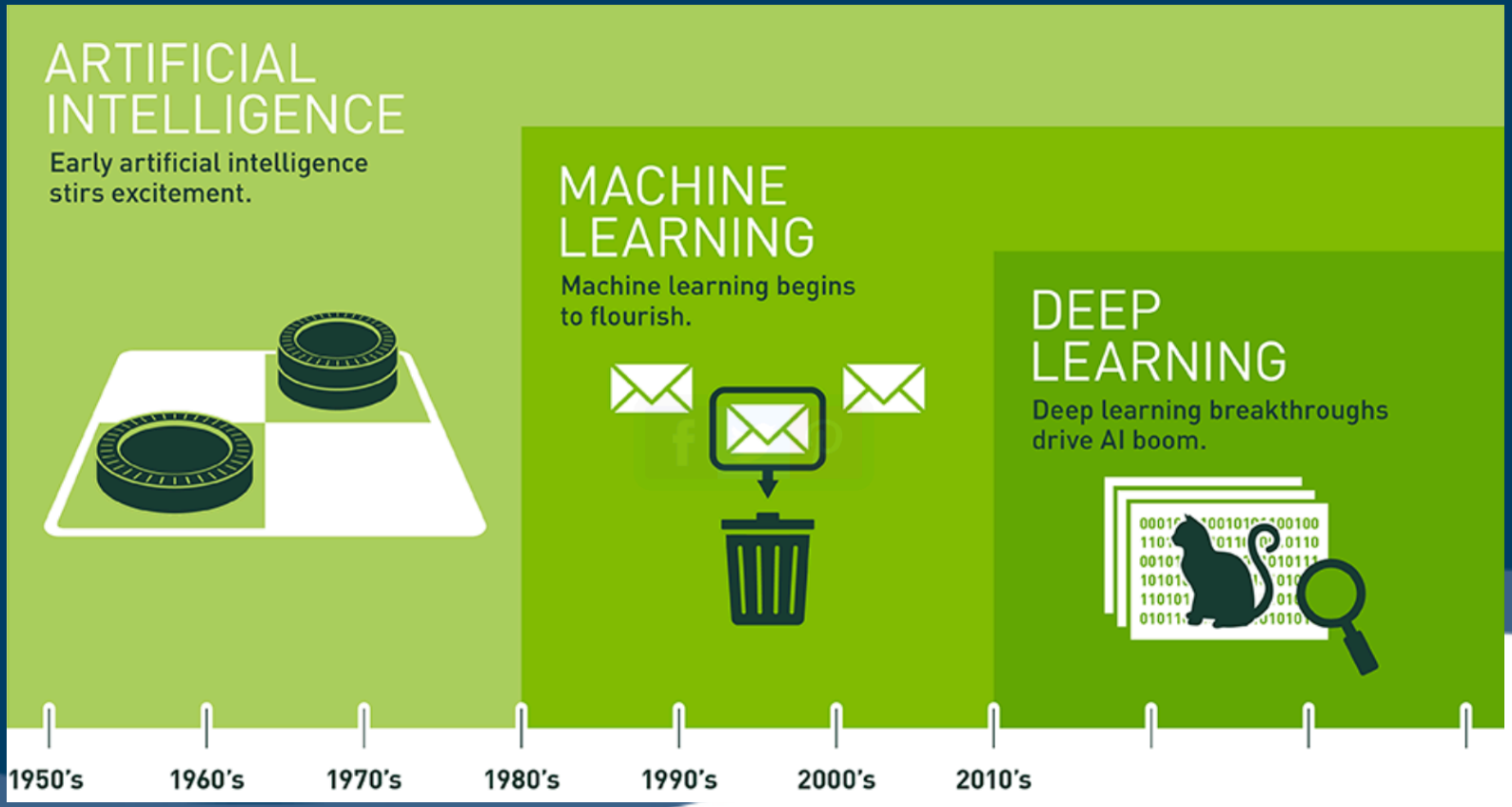
- Complexity:
  - Programming knowledge
  - Deep Learning
- Optimizing Deep Learning experience
- Analysis:
  - Data Augmentation code
- Pathologists

Objectives

- Identify CNN commonalities
- No programming expertise required

- No viable model ensured:
  - Parameters used influence model viability
  - Data preprocessing
  - Insight in CNN's still required

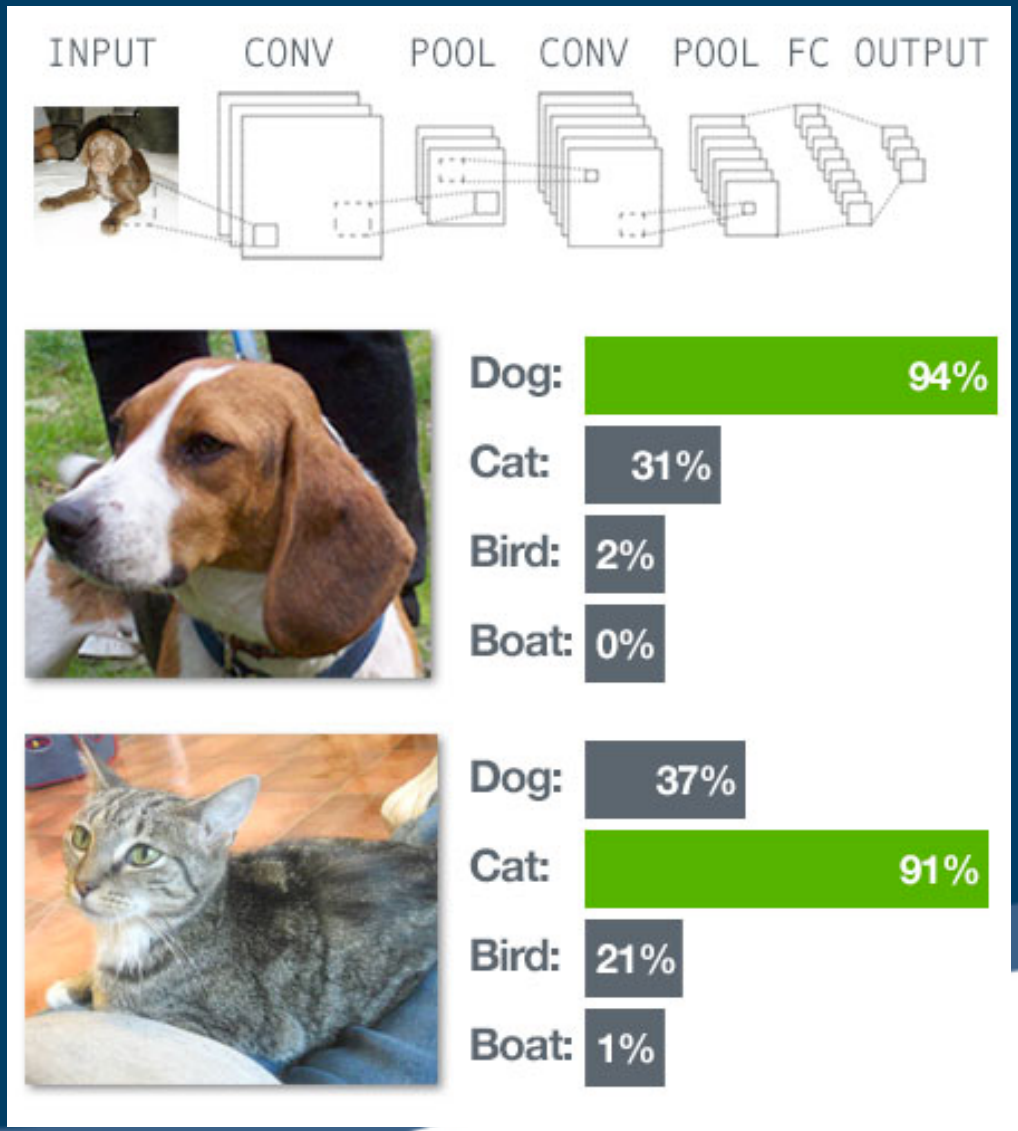
Machine Learning vs. Deep Learning



Universiteit Antwerpen

HistoGeneX

Bron: https://www.pinterest.cl/pin/705939310313420837/

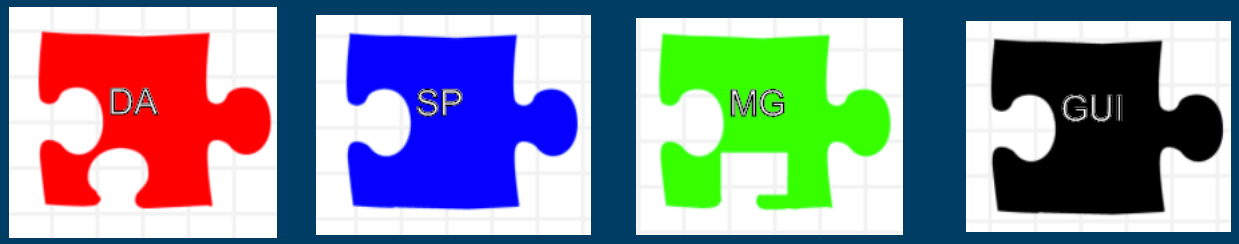


Universiteit Antwerpen

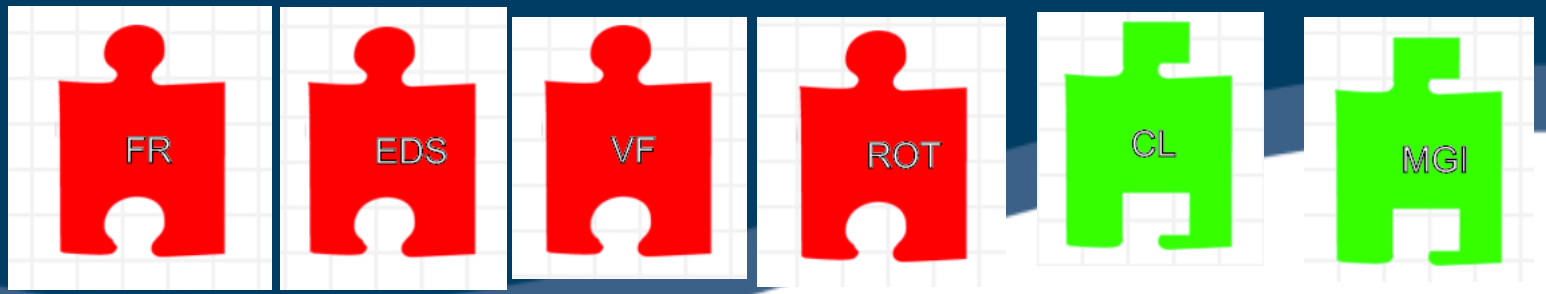
Bron: https://www.pyimagesearch.com/2016/06/27/my-top-9-favorite-python-deep-learning-libraries/

Puzzle Pieces

- Horizontal: Project Parts

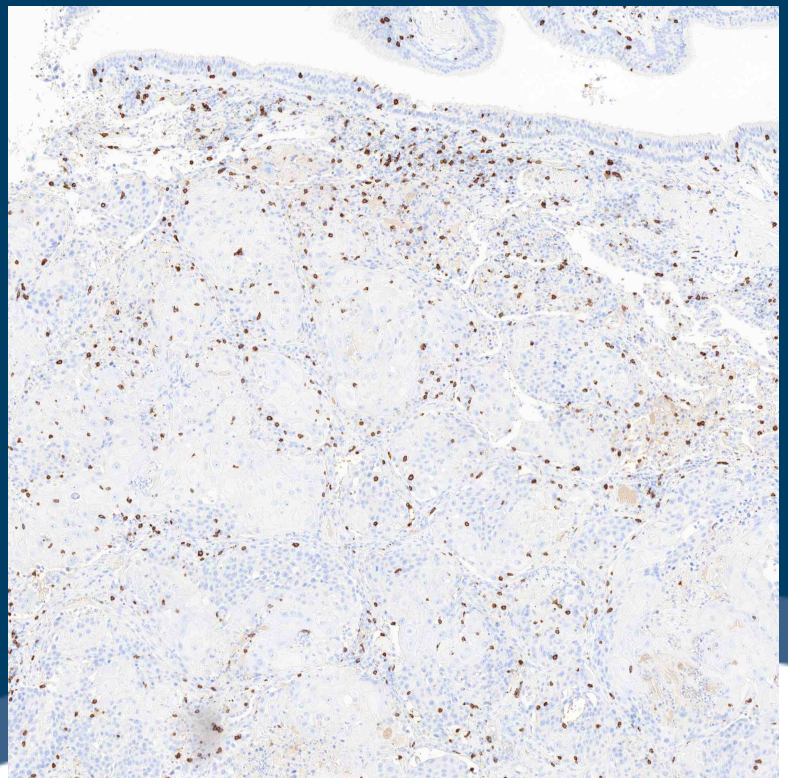


- Vertical: Specific Processes



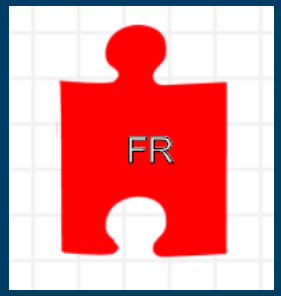
Data

- Images
- Labeled → Supervised Learning
- Specific folder hierarchy



Data Augmentation: Required Folder Hierarchy

- /data/
  - Dogs/
    - Dog.1.png
    - Dog.2.png
    - …
  - Cats/
    - Cat.1.png
    - Cat.2.png
    - …



Data Augmentation: Required Folder Hierarchy

- Renaming files → Automatable

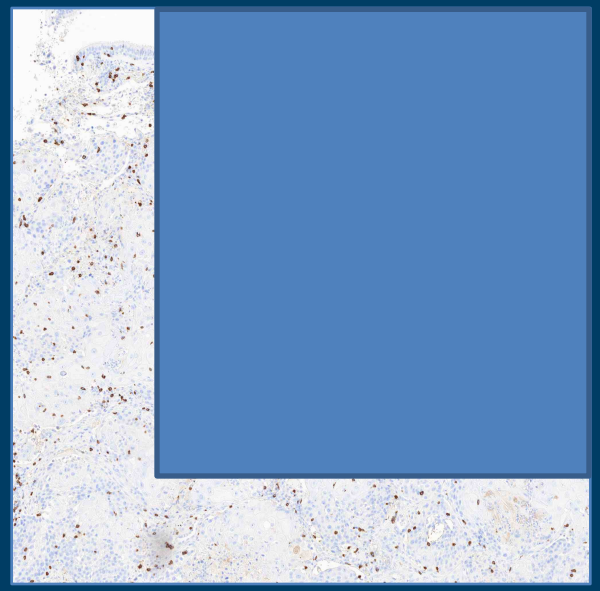
Data Augmentation: Expanding the dataset size

- Eg: 2000x2000 images
  - Take 1700x1700
  - Shift 10 pixels width/height
  - …
- 900 images from one image (30x30)

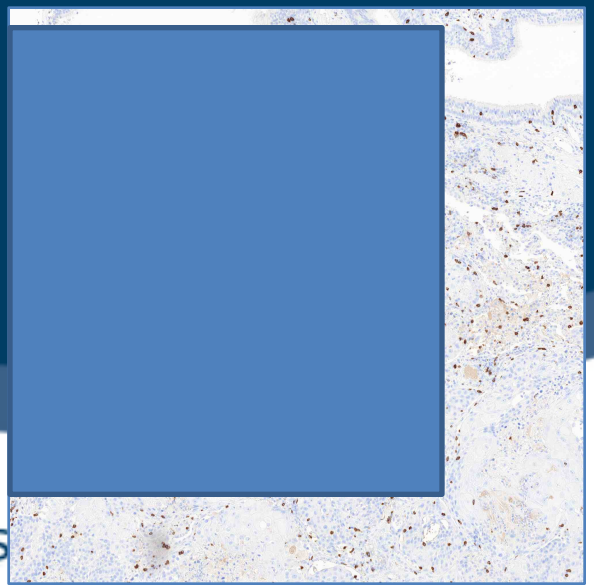
1

2

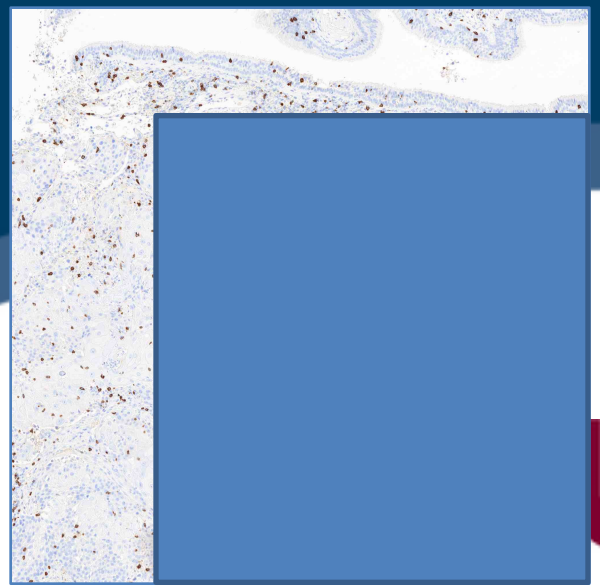
30



31

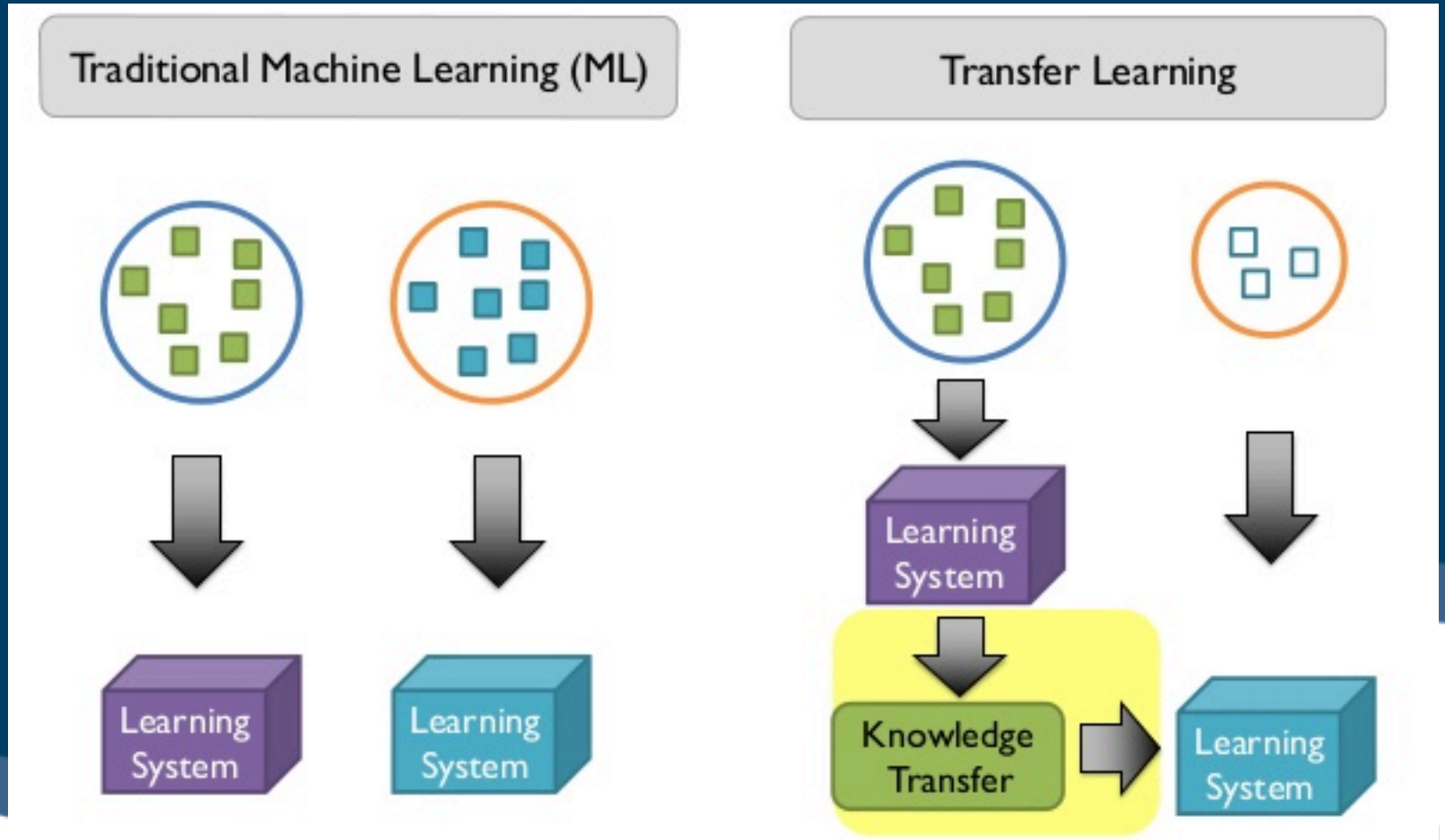


900



Univers

Transfer Learning & Fine Tuning

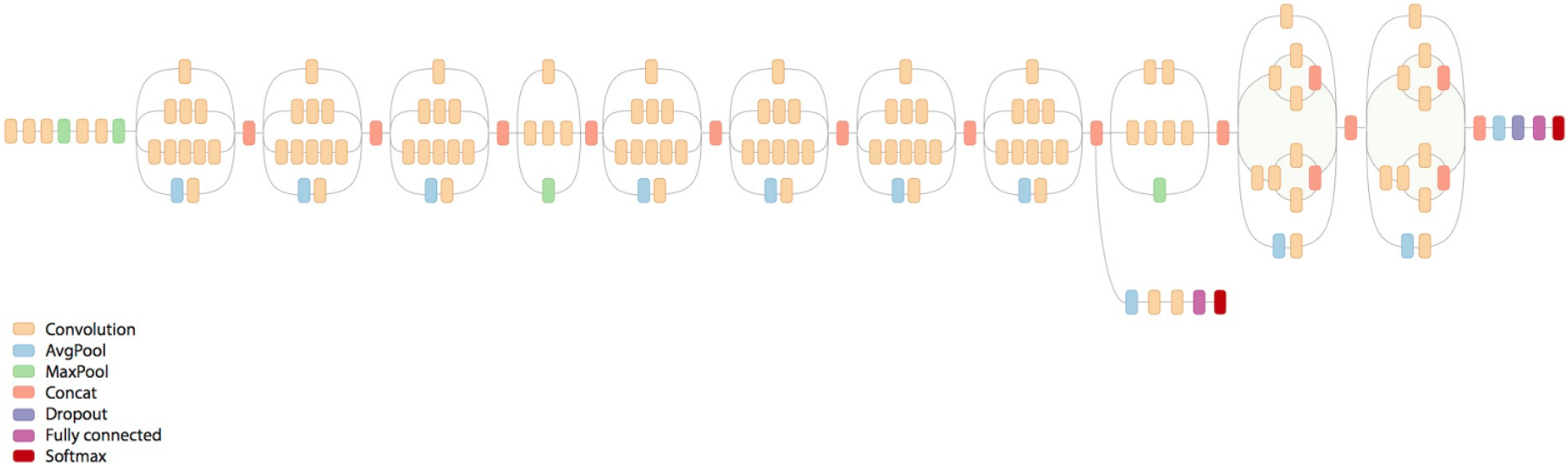


Universiteit Antwerpen

HistoGeneX

Bron: http://www.softfluent.fr/blog/expertise/2017/10/11/Compte-rendu-des-MS-Experiences-2017-Deep-Learning-choisir-son-framework-pour-en

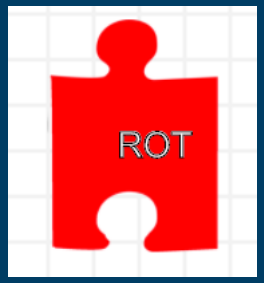
InceptionV3



Bron: https://research.googleblog.com/2016/03/train-your-own-image-classifier-with.html

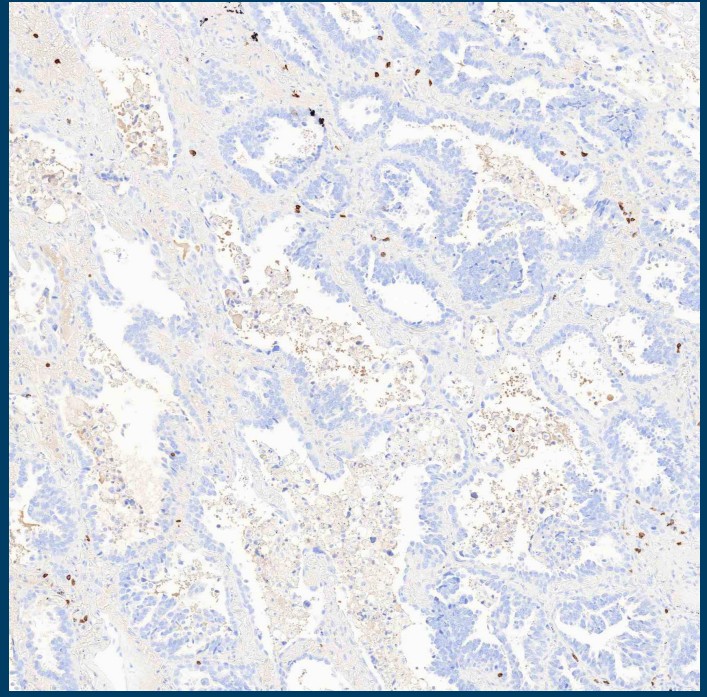
Data Augmentation: Expanding the dataset size

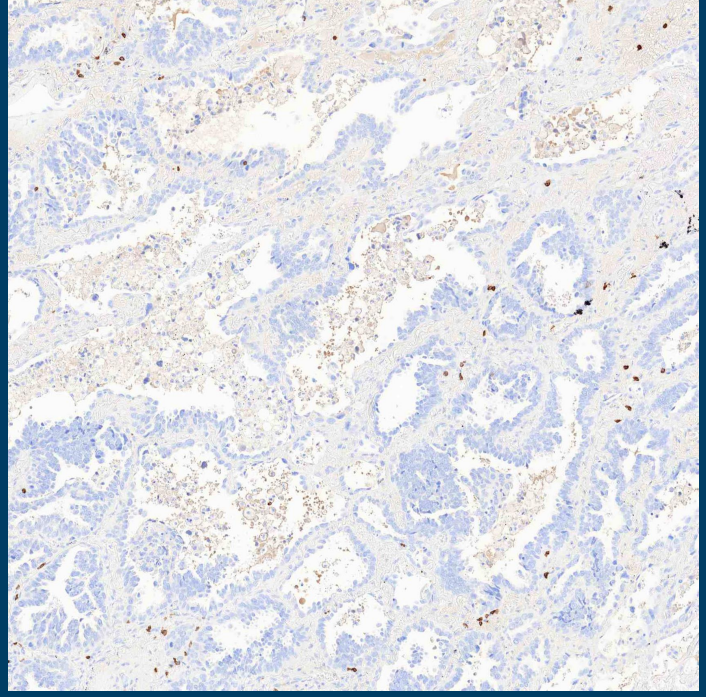
- Eg: 2000x2000 images
  - Take 1700x1700
  - Shift 10 pixels width/height
  - …
- 900 images from one image (30x30)
- InceptionV3 takes inputsize 299
  - Min width/height shift: 1700/299 < 6 pixels

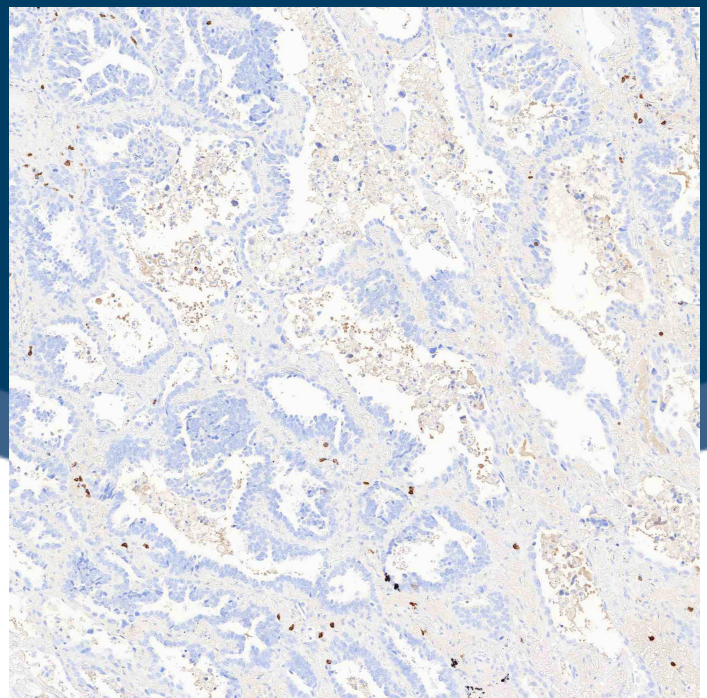


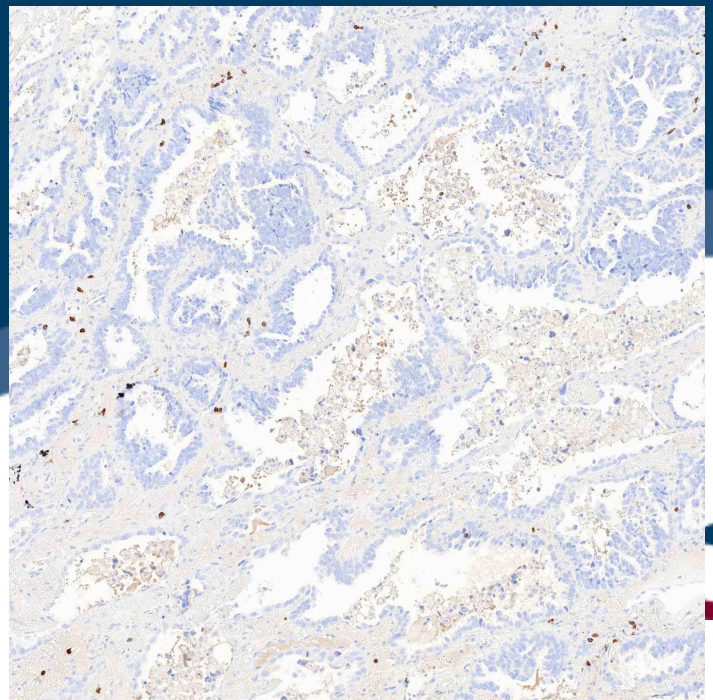
Data Augmentation: Expanding the dataset size

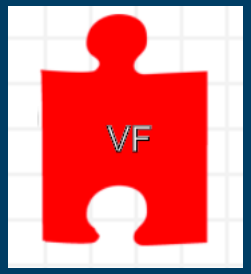
- Rotate 90 degrees:
  - 0
  - 90
  - 180
  - 270
- X4 images





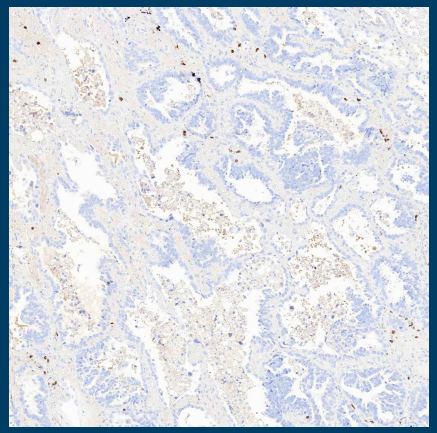
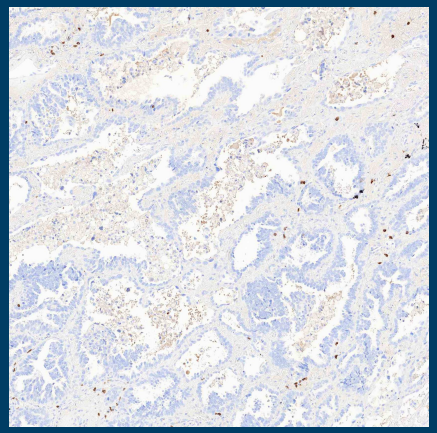
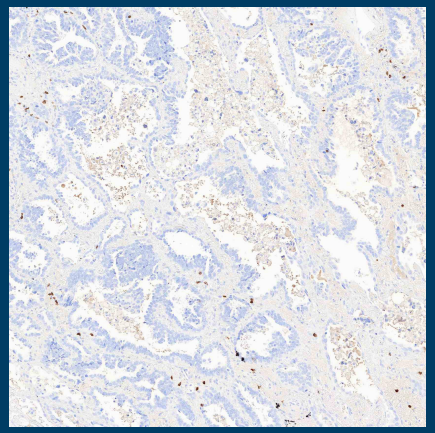
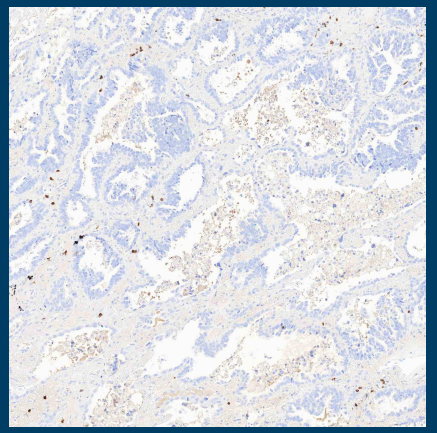
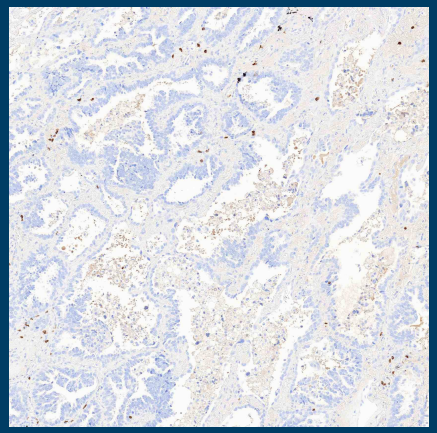
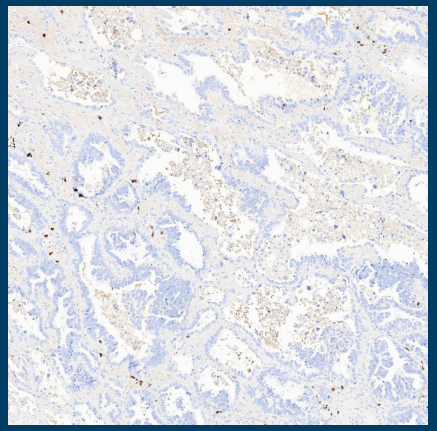
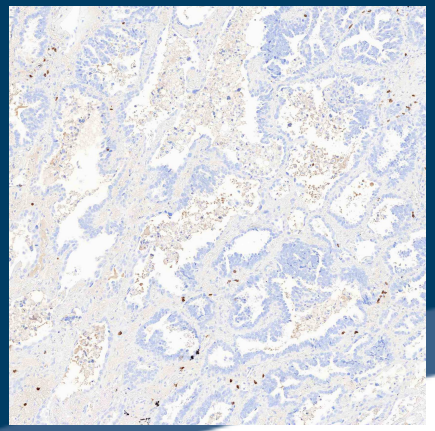
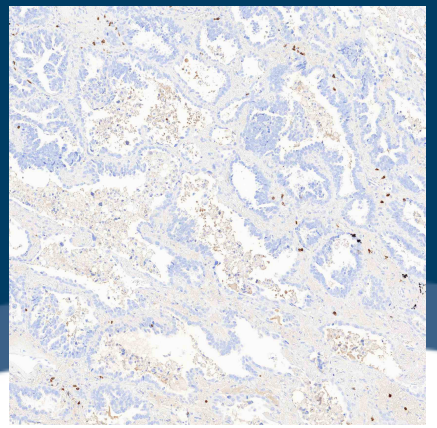






Data Augmentation: Expanding the dataset size

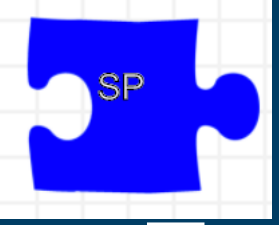
- Flip Vertically
- X2 images

Training, Validation and Testing set

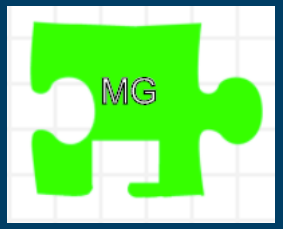
- 60% Training
- 20% Validation
- 20% Testing

- Specific resulting folder hierarchy



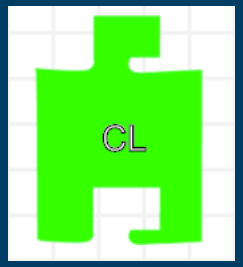
Training, Validation and Testing set

- /augmented_data/
  - Train_dir/
    - Dogs/
    - Cats/
  - Val_dir/
    - Dogs/
    - Cats/
  - Test_dir/
    - Dogs/
    - Cats/



Model generation

- Amount of epochs
- Transfer learning optimizers
- Optimizer learning rates
- Directory of:
  - Training data
  - Augmentation data

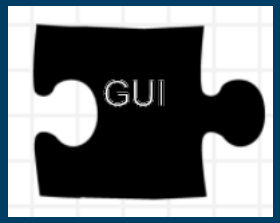


Classifiers

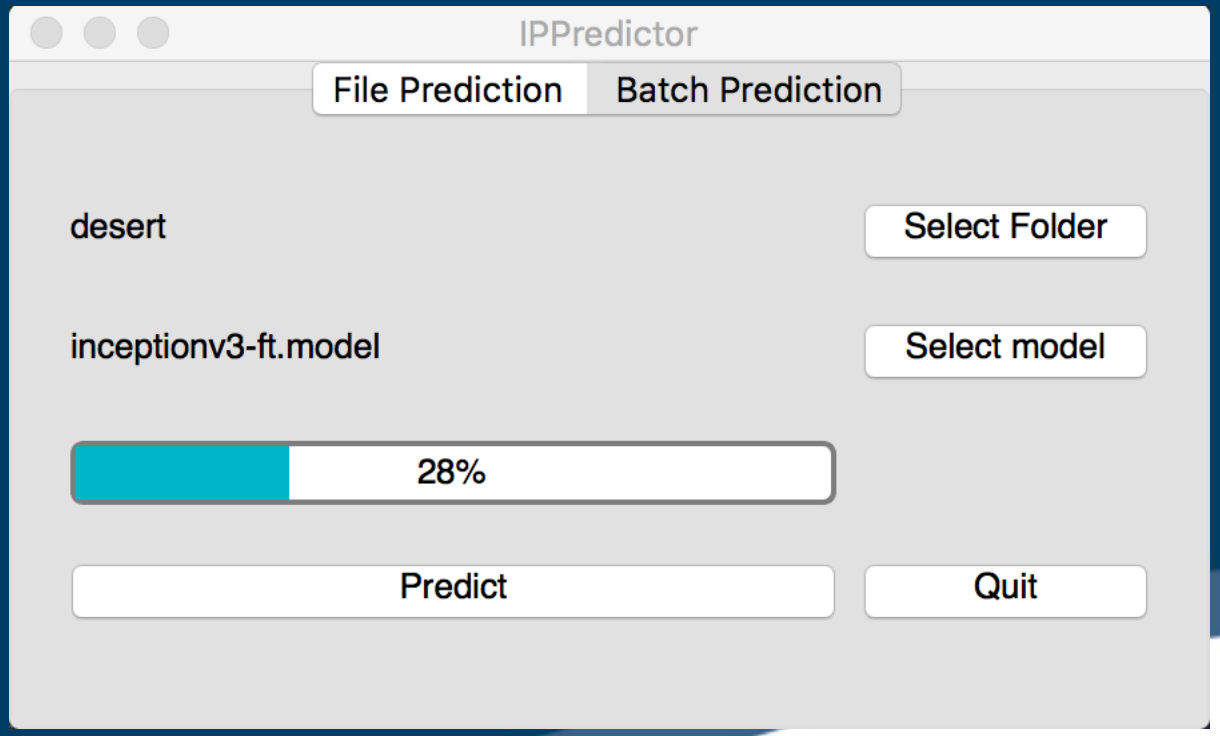
- Possible Classes
- Unbalanced Classes → Class weights
- Ex:
  - Class A Weight: 1.0
  - Class B Weight: 2.0
  - Class C Weight: 3.0

Model Generation Interruption

- Implements Callback
- Interruption on loss function
- Patience Parameter



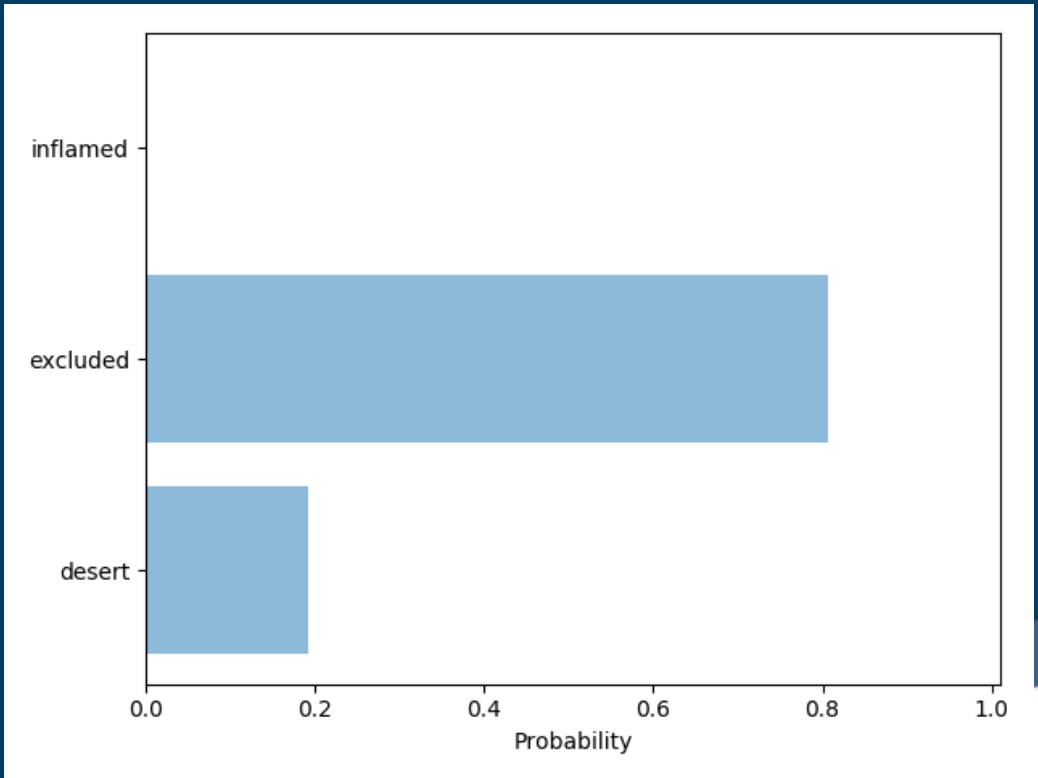
Gui



Universiteit Antwerpen

HistoGeneX

Prediction Plot Example



Universiteit Antwerpen

Tools

- AToMPM: Domain-specific modeling tools
- metaDepth: Template coding
- Python
- Keras: High-level neural networks API (TensorFlow, Theano)

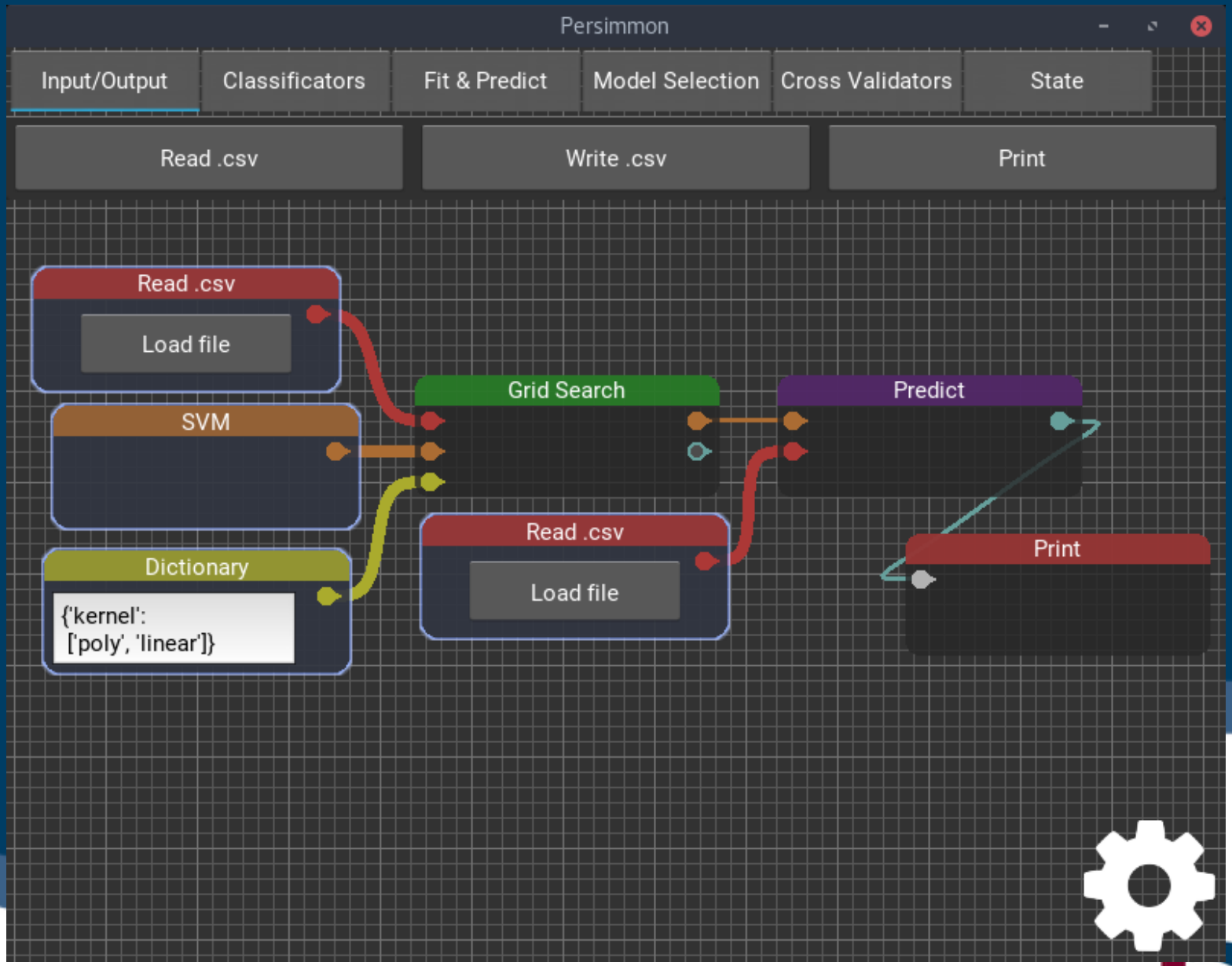
Demo

(https://cl.ly/3F0b470u1b0p)

Related Work: Persimmon

- Drag and drop scikit-learn tool for Machine Learning.
- Aids with the lack of programming skills
- Aim to provide the user with feasibility study functionality

- No functionality for CNN's



Universiteit Antwerpen

Bron: http://eprints.ucm.es/44618/1/Persimmon.pdf

Conclusion

- Generate CNN’s
- No Programming skills required
- Model can be tweaked using parameters
- Learning about CNN’s is required
- User Feedback