Using Deep Learning to Detect Galaxy Mergers

Jonas Arilho Levy

Instituto de Matemática e Estatística da Universidade de São Paulo

Objectives

- Detect galaxy mergers using Deep Learning techniques.
- Investigate 3 Convolutional Neural Networks (CNNs) architectures.
- Compare learning from scratch and transfer learning.
- Outperform previous automatic detection methods.

Background

- When two galaxies come close together, they start interacting with each other due to gravitational pull and this effect is called a galaxy merger.
- Astronomers use photometry to detect objects of interest and then fine-grained data is captured for those few selected using spectroscopy.
- Neural Networks were first introduced in 1943 [4], taking a biological inspiration by the way neural networks in animals and are used for supervised learning to predict whether a sample of the dataset belongs to a certain class or not.
- Gradient descent and backpropagation are used to minimize a cost function iteratively by moving in the direction of steepest descent to update the weights of the model.
- Convolutional Neural Networks use convolutions to transform an image by applying a kernel over each group of pixels in each layer of the network, with pooling and fully-connected layers to reduce dimensions of the image.

Architectures

1. **VGG-16** was created by members of the Visual Geometry Group (VGG) at the University of Oxford and is a deep network with 138 Million parameters, 13 convolutional layers and 3 fully-connected layers. [5]

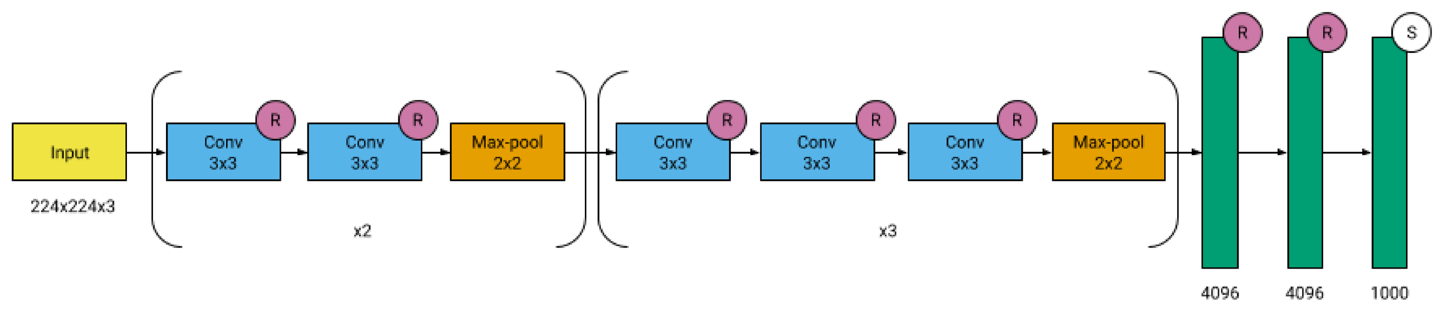

Figure 1: VGG-16 architecture

2. **Inception-v3** was inspired by the movie Inception and the quote *"We need to go deeper"* and stacks dense blocks of convolutional layers and uses batch normalisation in auxiliary layers, having a total of 24 million parameters. [6]

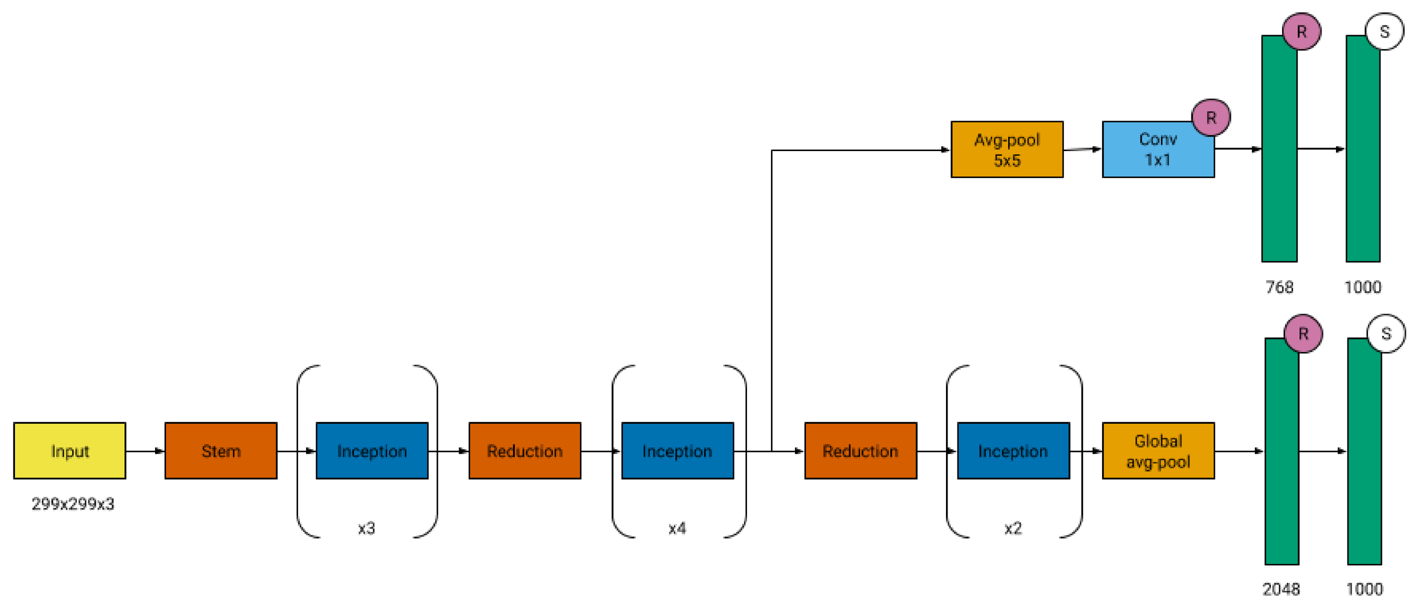

Figure 2: Inception-v3 architecture

3. **Densenet-121** adds shortcuts among layers, having only 0.8 Million trainable parameters because each layer receives feature maps from all preceding layers and also features a *growth rate* hyperparameter that defines how many additional channels are need for each layer. [2]

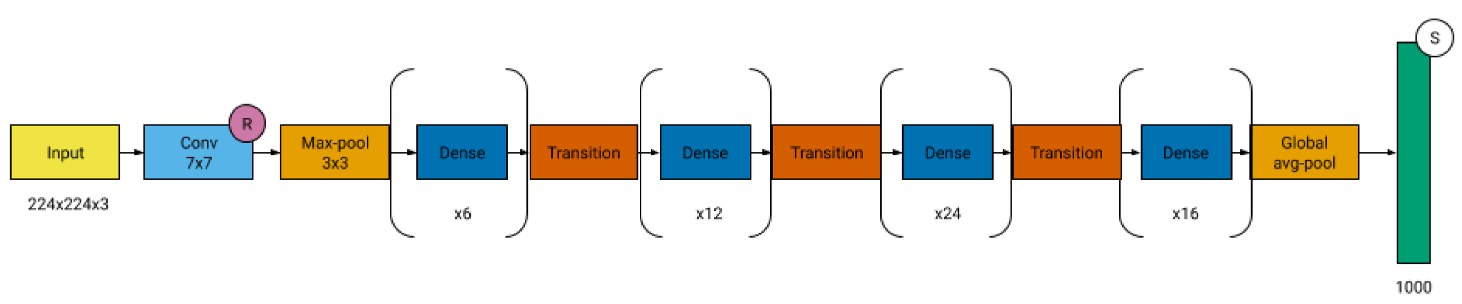

Figure 3: Densenet-121 architecture

Experiment 1

- Consists in training the 3 architectures from scratch, *i.e.*, initializing the model with random weights and training them using the training data, as follows:
  1. Use random initialization to the weights;
  2. Add top layers;
  3. Train using mini-batch SGD with a standard learning rate.

Experiment 2

- Consists in the fine tuning of pre-trained convolutional neural networks, *i.e.*, only the last convolutional blocks are trained alongside the top layers and the weights are loaded from a model pre-trained on the imagenet dataset [1]. This was done as follows:
  1. Load the pre-trained CNN with weights;
  2. Add top layers to the CNN and use the ADAM optimizer to train only the top layers;
  3. Fine-tune the last convolutional blocks using mini-batch SGD with a small learning rate and momentum.

Dataset

- The dataset used in this work is comprised of 16000 RGB images from the Sloan Digital Sky Survey (SDSS) Data Release 7.
- Contains two classes, namely *merger*, the one we are interested in detecting, and *non-interacting*, labeled by the Galaxy Zoo project. [3].

Figure 4: Galaxy mergers.

Figure 5: Non-interacting galaxies.

Results

- VGG-16 and Densenet-121 achieved a precision of 0.97 on this dataset using transfer learning .
- Most architectures had similar classification performance, and the main difference found was a slight improvement by using transfer learning.

Table 1: Precision (p), Recall(r) and F1-score(F_1) for each architecture and experiment (E).

Architecture	E	p	r	F_1
VGG-16	1	0.96	0.96	0.96
	2	0.97	0.97	0.97
Inception-V3	1	0.96	0.96	0.96
	2	0.25	0.37	0.20
Densenet-121	1	0.96	0.96	0.96
	2	0.97	0.97	0.97

Conclusions

- A high accuracy can be achieved by using multiple Deep Learning techniques and architectures in this dataset.
- By using transfer learning from the *Imagenet* dataset to the images of merging galaxies there was a slight increase in performance.
- This study presents a reliable approach to detect galaxy mergers that outperforms previous automatic detection methods .

Bibliography

[1] J. Deng et al. "ImageNet: A Large-Scale Hierarchical Image Database". In: *CVPR09*. 2009.

[2] Gao Huang et al. "Densely connected convolutional networks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 4700–4708.

[3] Chris Lintott et al. "Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies*". In: *Monthly Notices of the Royal Astronomical Society* 410.1 (Dec. 2010), pp. 166–178. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2010.17432.x.

[4] Warren S McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *The bulletin of mathematical biophysics* 5.4 (1943), pp. 115–133.

[5] Karen Simonyan and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition". In: *arXiv preprint arXiv:1409.1556* (2014).

[6] Christian Szegedy et al. "Rethinking the inception architecture for computer vision". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2818–2826.

Monograph and code available at https://linux.ime.usp.br/~jonasarilho/mac0499/