# Classification of Histopathological Lung Cell Cancers Images using Deep Learning

## Project Report

Sahil Khose[1] Ankita Ghosh[2]

[1]3rd year, ICT Department
Manipal Institute of Technology

[2]3rd year, CSE Department
Manipal Institute of Technology

4 June 2021

# Table of Contents

# Histopathological Data Challenges

- The data slides captured are gigabytes in size and thus pose a computation bottleneck both in terms of storage, processing and compatibility with deep learning algorithms.
- Common approaches to tackle the problem:
  - patch extraction/ tiling method
  - downsampling to a tractable size
- In case we have a smaller dataset, we combined semi, weakly and un-supervised training techniques

# Learning Approaches

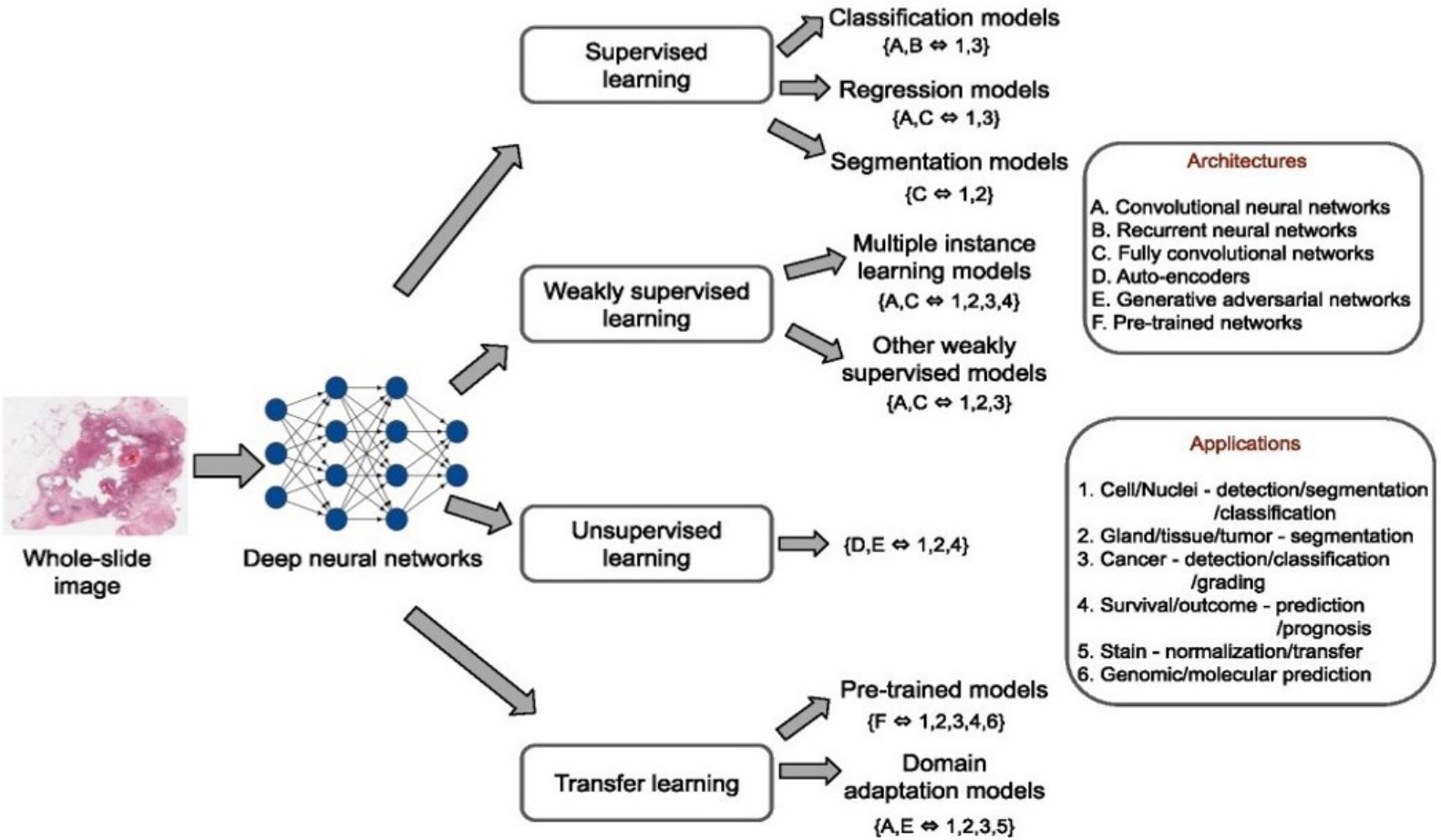
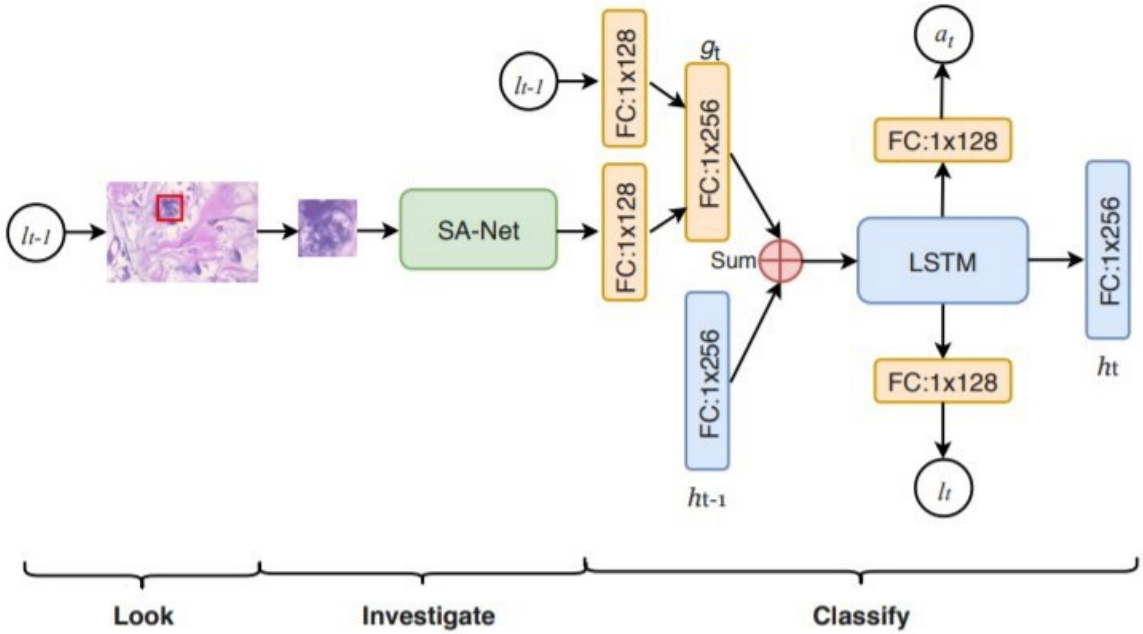



Image from Srinidhi et al.

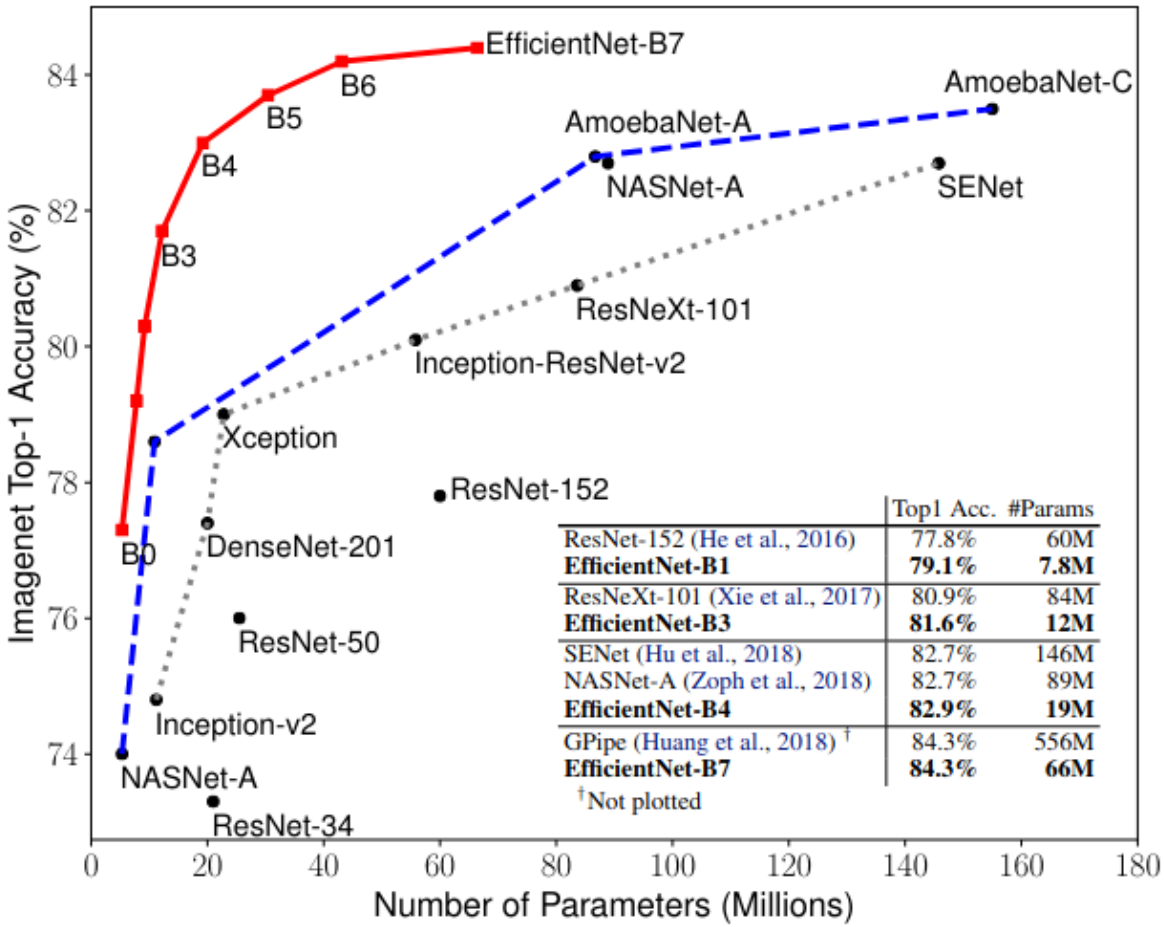
# Xu et al.

- They adaptively select a sequence of coarse regions by a hard visual attention algorithm, for each such region, it is able to investigate the abnormal parts based on a soft-attention mechanism.
- A recurrent network is then built on top to classify the image region and also to predict the location of the image region to be investigated at the next time step.



# Six-type cancer classifier

EfficientNet-B5 model seemed to outperform ResNet-50 especially when the slides were collected from multiple sources which requires learning abstract features that help in distinguishing the different classes.

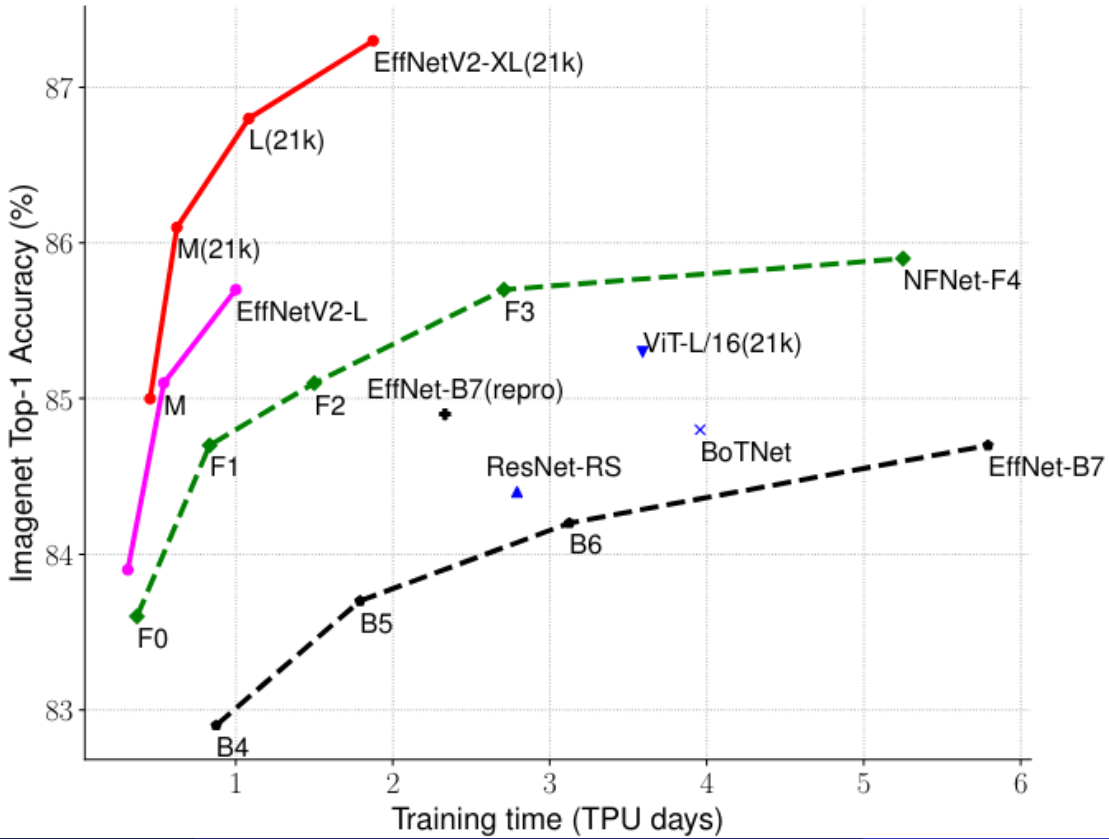



| | Top1 Acc. | #Params |
|---|---|---|
| ResNet-152 (He et al., 2016) | 77.8% | 60M |
| **EfficientNet-B1** | **79.1%** | **7.8M** |
| ResNeXt-101 (Xie et al., 2017) | 80.9% | 84M |
| **EfficientNet-B3** | **81.6%** | **12M** |
| SENet (Hu et al., 2018) | 82.7% | 146M |
| NASNet-A (Zoph et al., 2018) | 82.7% | 89M |
| **EfficientNet-B4** | **82.9%** | **19M** |
| GPipe (Huang et al., 2018) † | 84.3% | 556M |
| **EfficientNet-B7** | **84.3%** | **66M** |

†Not plotted

# EfficientNet V2

- The authors had access to only EfficientNet-B5 model implementation and model weights when the paper was implemented.
- We have access to B6-B8 models along with better architectures like NFNets, ViT, and EfficientNet V2 which was realeased 2 weeks ago



# Analysis Report of 130 Histopathology Papers

- They observe an exponentially growing trend for the number of papers published in this area.

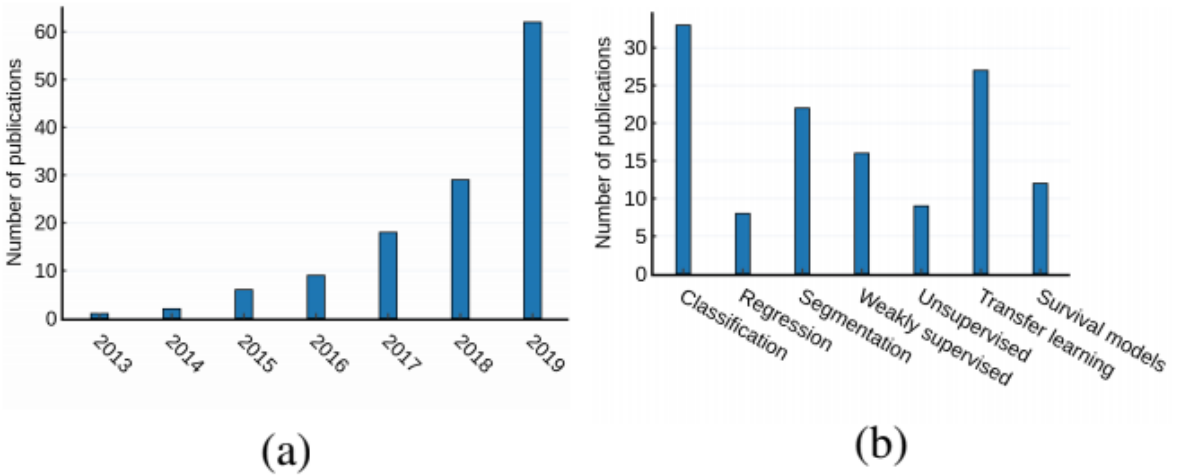



Fig. 1: (a) An overview of numbers of papers published from January 2013 to December 2019 in deep learning based computation histopathology surveyed in this paper. (b) A categorical breakdown of the number of papers published in each learning schemas.

# Analysis Report of 130 Histopathology Papers

- They compared the cancer types, staining, application, method and dataset of the abundance of papers they have collected.
- In most applications, standard architectures like VGGNet, InceptionNet, ResNet, MobileNet, DenseNet can be directly employed, and custom networks should only be used if it is impossible to transform the inputs into a suitable format for the given architecture, as the transformation may cause significant information loss that may affect the task performance.
- Given the dataset with few images, training only the last decision layers while freezing the initial layers, using a non-linear decision layer, using regularization techniques like weight decay and dropout are recommended to avoid overfitting.

# Concluding Discussion

Our primary concern at the moment is to search and retrieve a dataset which suits our problem statement.

As this problem statement has been approached by implementing some of the best models like EfficientNet, we are looking forward to more innovation in this project alongside implementing the best performing architectures.