Building a Convolutional Network for Brain Tumor Segmentation

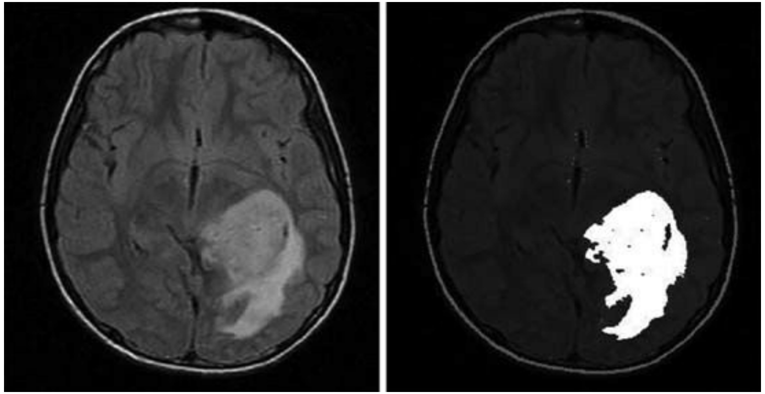
Intelligent systems for pattern recognition (760AA)

Giovanni Scognamiglio

What we did

- **Research**
  - understanding the assumptions behind convolutional architectures & modules
  - identifying lightweight module innovations
- **Model design and implementation**
  - choosing the right architecture and modules for the task at hand
- **Test the final model**

Introduction



- **What is brain tumor segmentation?**
  - brain tumor and MRIs
  - benefits of automatically segmenting tumor area in MRIs
- **Why are neural network ideal for identifying brain tumor?**
  - fuzzy borders, hard to distinguish
  - no priors about location, shape or contrast

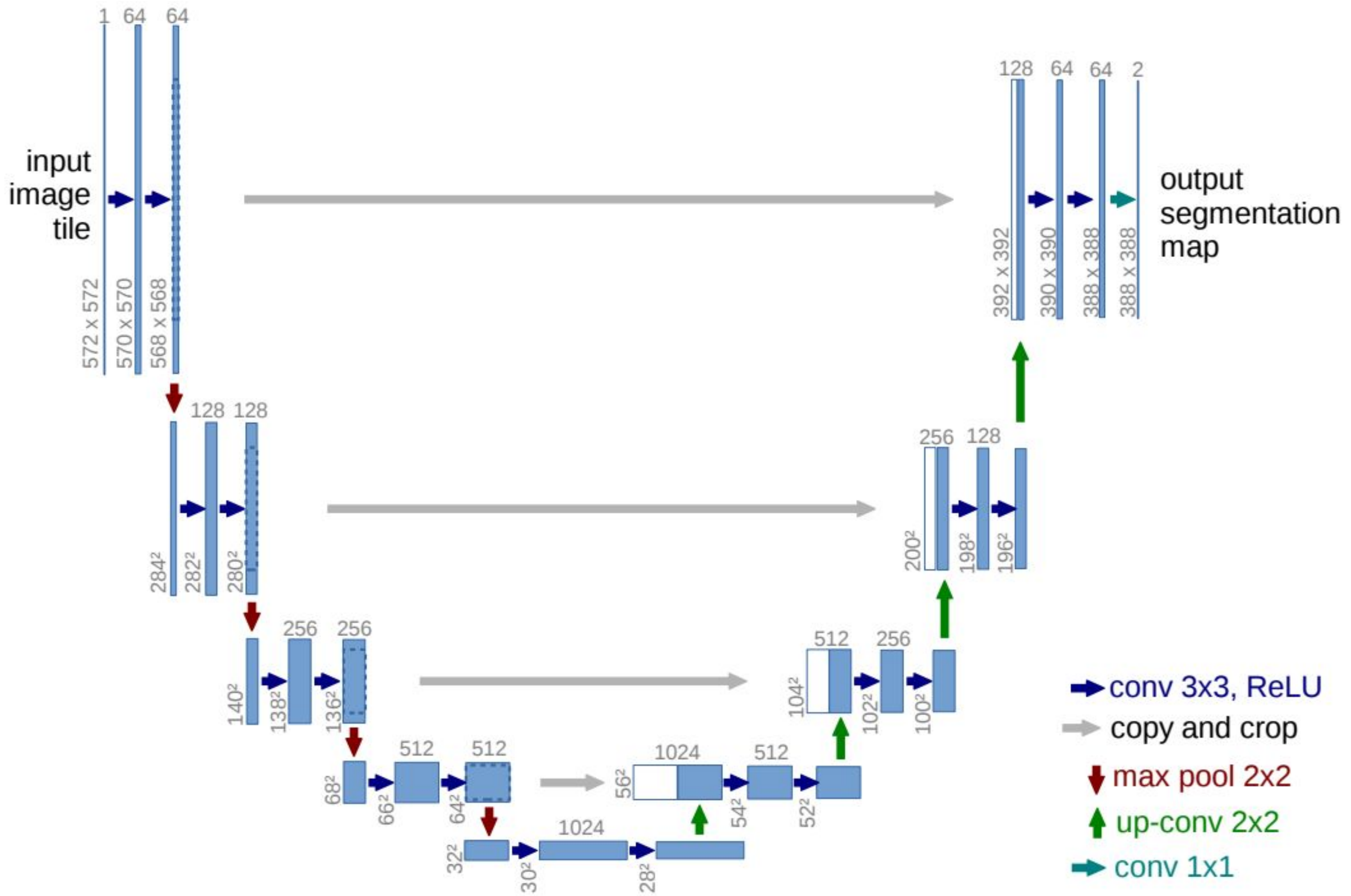
Introduction

- **The dataset we used**
  - created by researchers to investigate statistical relationship between tumor shape and genomic data
  - available on kaggle with 200+ notebooks with deep CNN (90% implementations are U-Net)
- **Limitations and opportunities**
  - Sporadic access to Colab GPU meant we had to take a more theoretical approach for model selection

Research - segmentation models

- **Understanding U-Net popularity**
  - what is U-Net
  - the pillars of U-Net
    - 3x3 conv
    - encoder - decoder structure
    - convolutional transpose
  - why are U-Net modules/architecture effective for medical image segmentation
  - what came after U-Net

U-Net

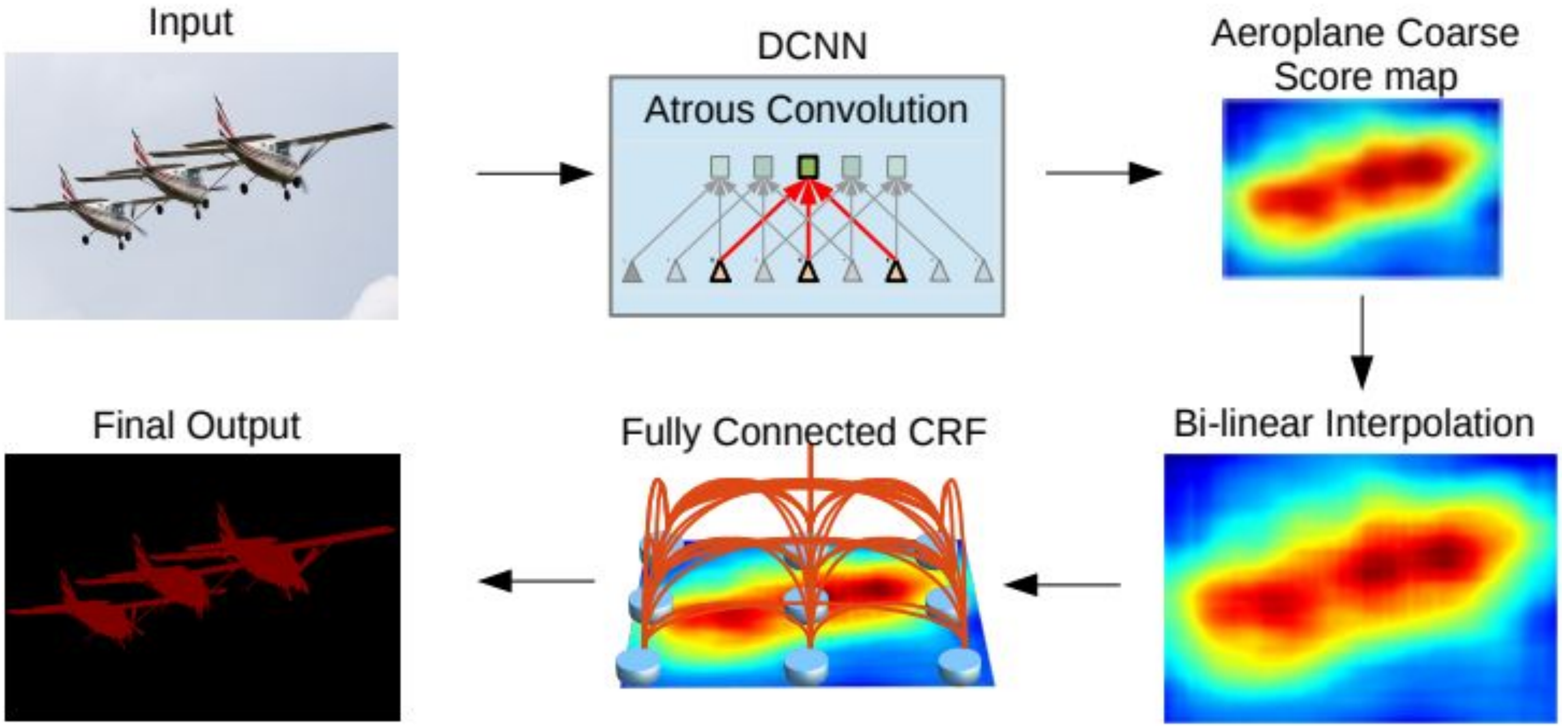


1 64 64
input image tile
572 x 572
570 x 570
568 x 568
128 128
284^2
282^2
280^2
256 256
140^2
138^2
136^2
512 512
68^2
66^2
64^2
1024
32^2
30^2
28^2
1024
512
56^2
54^2
52^2
512 256
104^2
102^2
100^2
256 128
200^2
198^2
196^2
128 64 64 2
392 x 392
390 x 390
388 x 388
388 x 388
output segmentation map
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1

Research - segmentation models

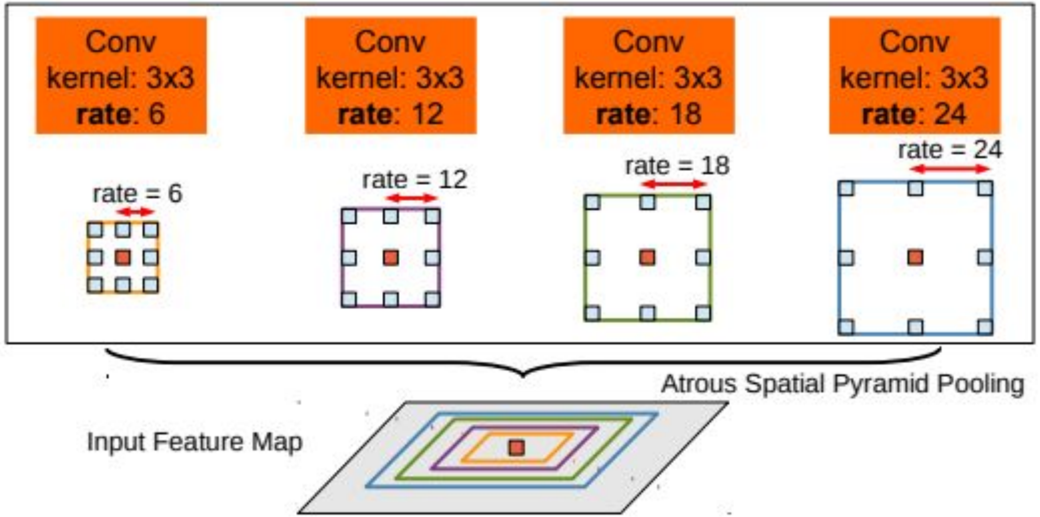
- **Understanding DeepLab: U-Net’s antagonist**
  - is encoding actually necessary for image segmentation?
  - what is “multi-scale contextual learning”?
  - the pillars of DeepLab:
    - dilated convolutions (a double edged sword)
    - Atrous Spatial Pyramid Pooling (ASPP)
- a note on dilated convolutions: current limitations and future possibilities

DeepLab


Input
DCNN
Atrous Convolution
Aeroplane Coarse
Score map
Bi-linear Interpolation
Fully Connected CRF
Final Output

Atrous Spatial Pooling (ASPP)

(we will use it later)



Research - lightweight models

- **Why researching lighter models?**
- **Our focus: parameter-efficient models.**
    - reduction in # of parameters and multiplications-additions (mult.adds)
- How?
    - Factorizing convolutions
    - reducing redundancy

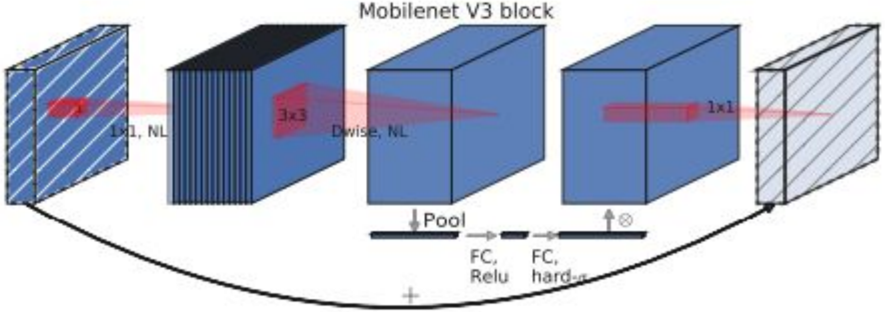
Research - lightweight models

- **Factorizing convolutions**
    - **Depthwise separable convolutions**
    - Group convolutions
    - many other ways...
- Reducing redundancy
    - Residual connections
    - Dense feed-forward connections
    - many other ways...

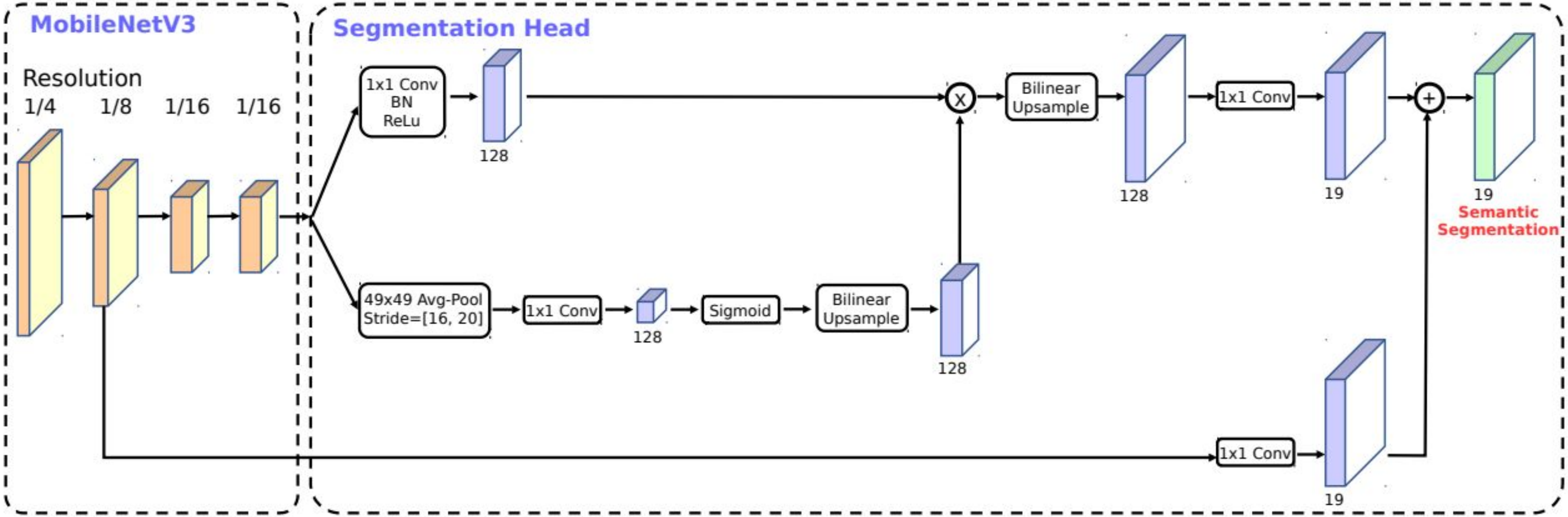
Model implementation: MobileNetv3

- Why we chose MobileNetv3
    - 1: research and papers
    - 2: code and documentation
    - 3: pre-designed segmentation head

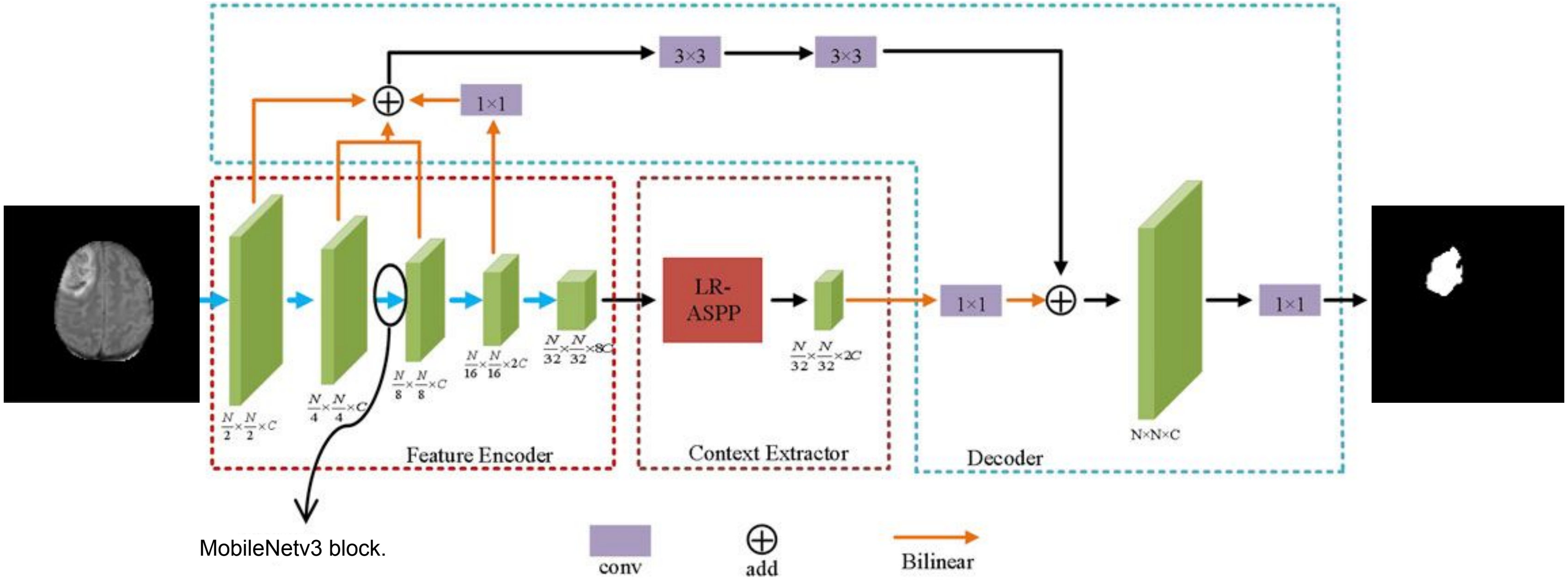
Input	Operator	exp size	$\#out$	SE	NL	s
$224^2 \times 3$	conv2d, 3x3	-	16	-	HS	2
$112^2 \times 16$	bneck, 3x3	16	16	✓	RE	2
$56^2 \times 16$	bneck, 3x3	72	24	-	RE	2
$28^2 \times 24$	bneck, 3x3	88	24	-	RE	1
$28^2 \times 24$	bneck, 5x5	96	40	✓	HS	2
$14^2 \times 40$	bneck, 5x5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5x5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5x5	120	48	✓	HS	1
$14^2 \times 48$	bneck, 5x5	144	48	✓	HS	1
$14^2 \times 48$	bneck, 5x5	288	96	✓	HS	2
$7^2 \times 96$	bneck, 5x5	576	96	✓	HS	1
$7^2 \times 96$	bneck, 5x5	576	96	✓	HS	1
$7^2 \times 96$	conv2d, 1x1	-	576	✓	HS	1
$7^2 \times 576$	pool, 7x7	-	-	-	-	1
$1^2 \times 576$	conv2d 1x1, NBN	-	1024	-	HS	1
$1^2 \times 1024$	conv2d 1x1, NBN	-	k	-	-	1



MobileNetv3



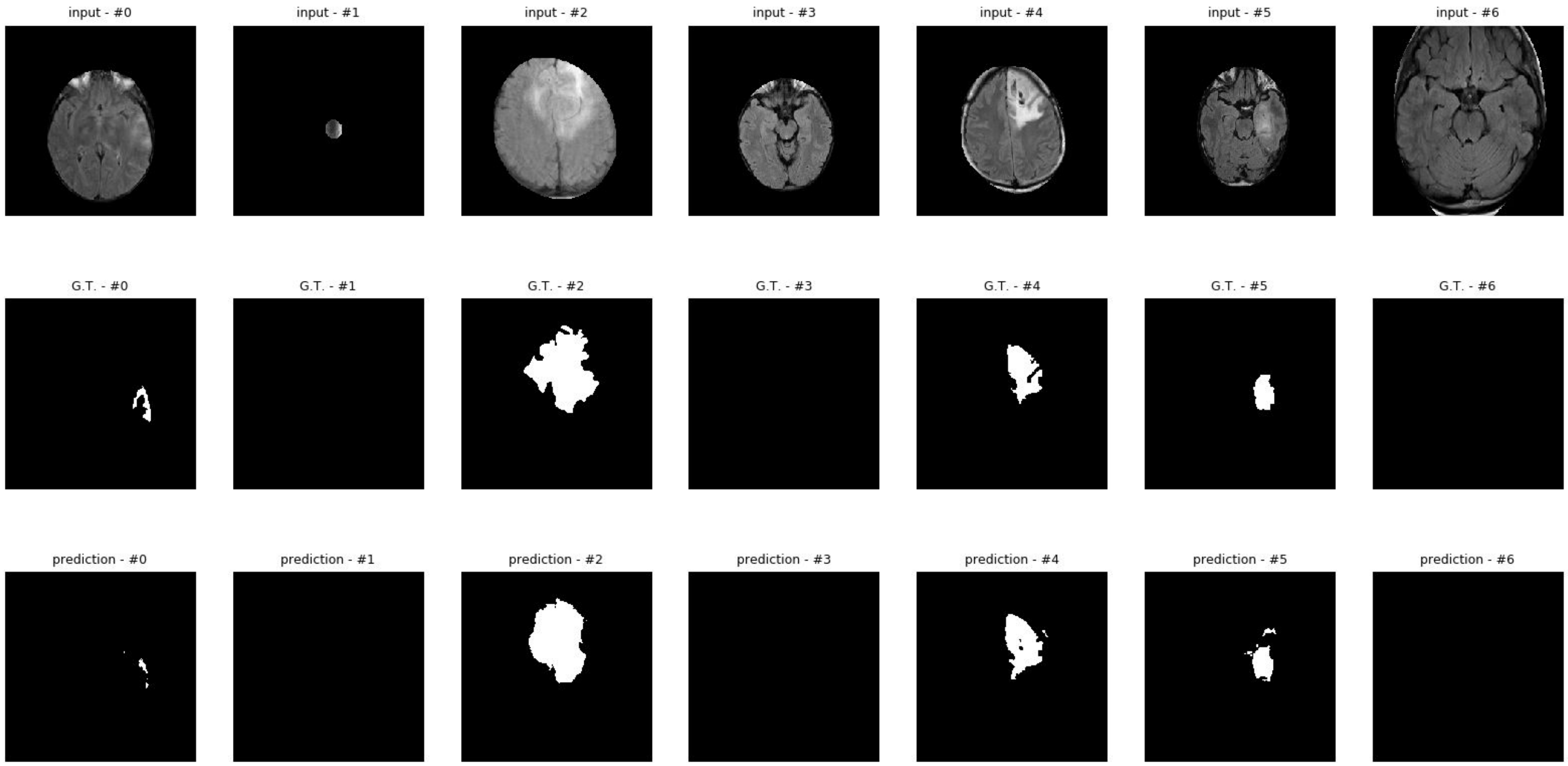
Hybrid MobileNetv3 + AL-NET



Implementation and Results

- Model selection for two models:
  - MobileNetV3
  - MobileNetV3 with extra skip connections

2*Models	2*# of trainable parameters	Hyperparpameters					DICE coef	
		opt	lr	mom	nest	bn	DICE VAL	**DICE TEST**
MobileNetv3 + LR-ASPP	310 thousands	SGD	0.09	0.95	Yes	Yes	83.49%	**84.76%**
MobileNetv3 + LR-ASPP with extra skip connections	312 thousands	SGD	0.09	0.95	Yes	Yes	84.01%	**85.26%**
U-Net (pre-trained)	7 milion	-	-	-	-	-	71.12%	**74.12%**
U-Net (claimed)	7 milion	-	-	-	-	-	-	**82%**


input - #0
input - #1
input - #2
input - #3
input - #4
input - #5
input - #6
G.T. - #0
G.T. - #1
G.T. - #2
G.T. - #3
G.T. - #4
G.T. - #5
G.T. - #6
prediction - #0
prediction - #1
prediction - #2
prediction - #3
prediction - #4
prediction - #5
prediction - #6

Conclusion

1. Consider lightweight models before trying U-Net as default go to
2. A need for general-purpose lightweight architectures
3. MobileNetv3 proved it self as a solid lightweight general purpose architecture