Project B
Facial Emotion Recognition

ECE1512 Digital Image Processing and Applications

Department of Electrical and Computer Engineering
University of Toronto

Winter 2021

Karush Suri

Logistics

- Project B uploaded
- Deadline-Monday, April 5, 5:00 PM EST.
- Please start early!
- Questions..
  - FAQs page (**primary source**)
  - Discussion board
  - TA by email (“Project-B” as subject line)

Today

- Tutorial (~ 1.5 hours)
  - Part-1: Facial Expression Recognition
  - Part-2: Facial Action Units
  - Part-3: Overview of PyTorch
- Break (~ 5 minutes)
- Open question hour (~ 0.5 hour)
  - Questions, questions, questions…

In case you have a question-

- Raise hand and unmute Mic
- Type in Chat

Today

- Tutorial (~ 1.5 hours)
  - Part-1: Facial Expression Recognition
  - Part-2: Facial Action Units
  - Part-3: Overview of PyTorch
- Break (~ 5 minutes)
- Open question hour (~ 0.5 hour)
  - Questions, questions, questions…

In case you have a question-

- Raise hand and unmute Mic
- Type in Chat

Today

- Tutorial (~ 1.5 hours)
  - Part-1: Facial Expression Recognition
  - Part-2: Facial Action Units
  - Part-3: Overview of PyTorch
- Break (~ 5 minutes)
- Open question hour (~ 0.5 hour)
  - Questions, questions, questions…

In case you have a question-

- Raise hand and unmute Mic
- Type in Chat

Facial Expression Recognition

What is an Expression?

the process of making known one's thoughts or feelings

Facial Expression Recognition

~~**What is an Expression?**~~

What is a Facial Expression?

a look on someone's face that conveys a particular emotion

Facial Expression Recognition

- Humans mostly communicate expressions via face
- Positions and movement of facial components express our feeling and emotions
- Often more powerful and expressive in comparison to verbal communication
- Ease of interpretation via human perception

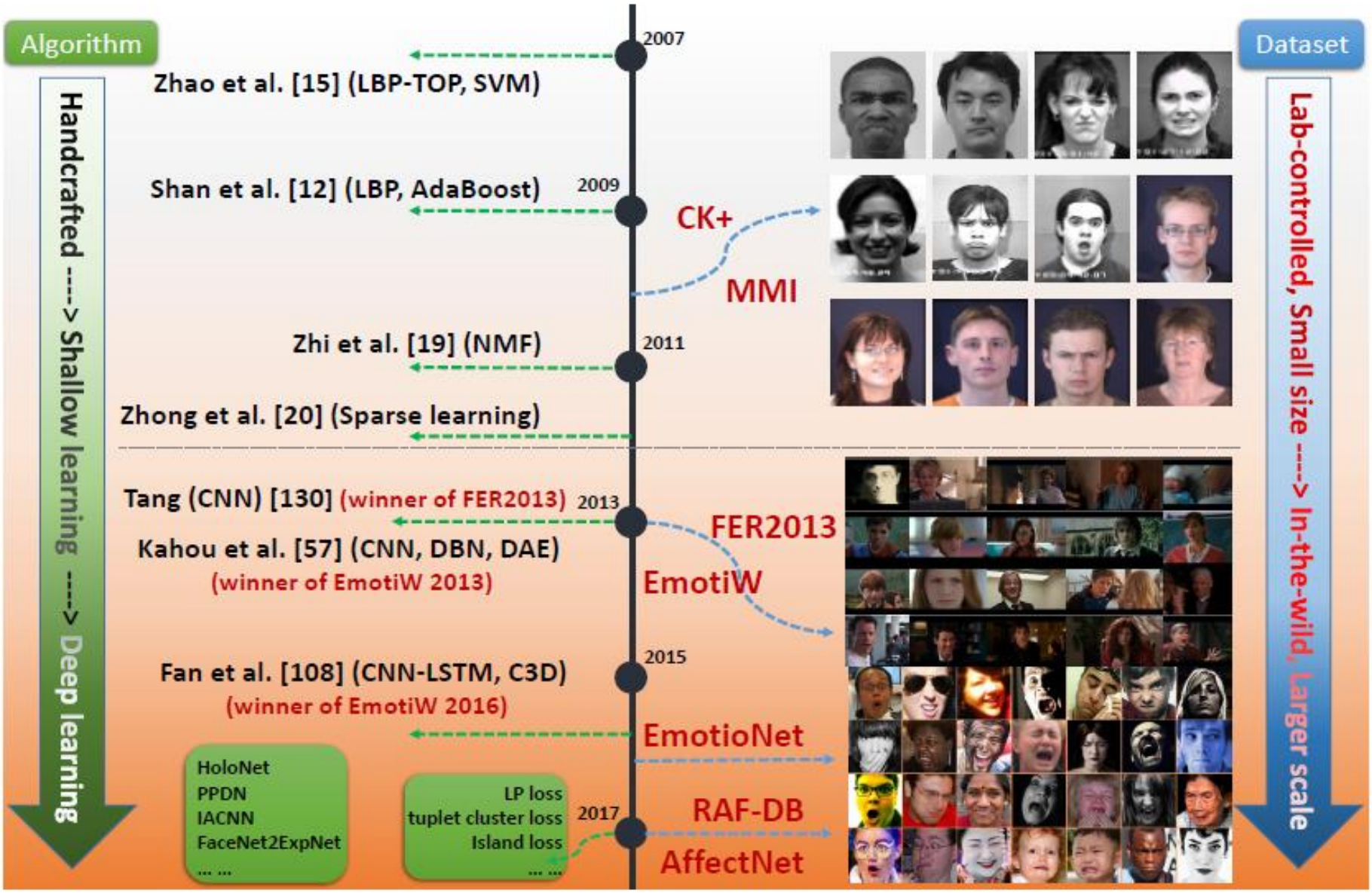
Facial Expression Recognition

We are interested in automatic perception of expressions

Drum roll…

Enter the era of Machine Learning

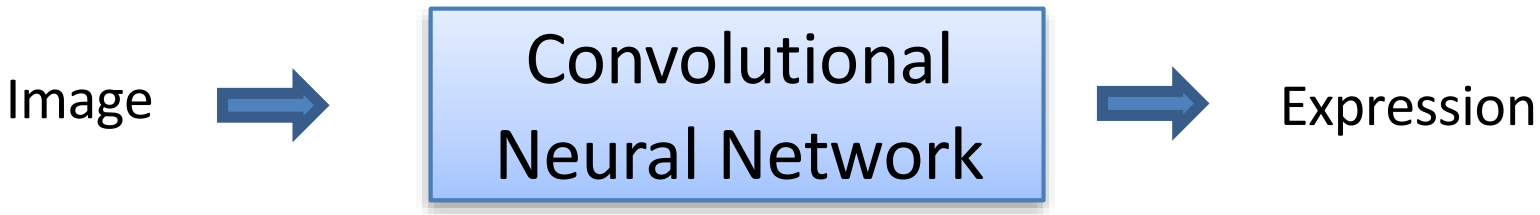
Facial Expression Recognition



Shan Li and Weihong Deng, *Deep Facial Expression Recognition: A Survey*, 2018.

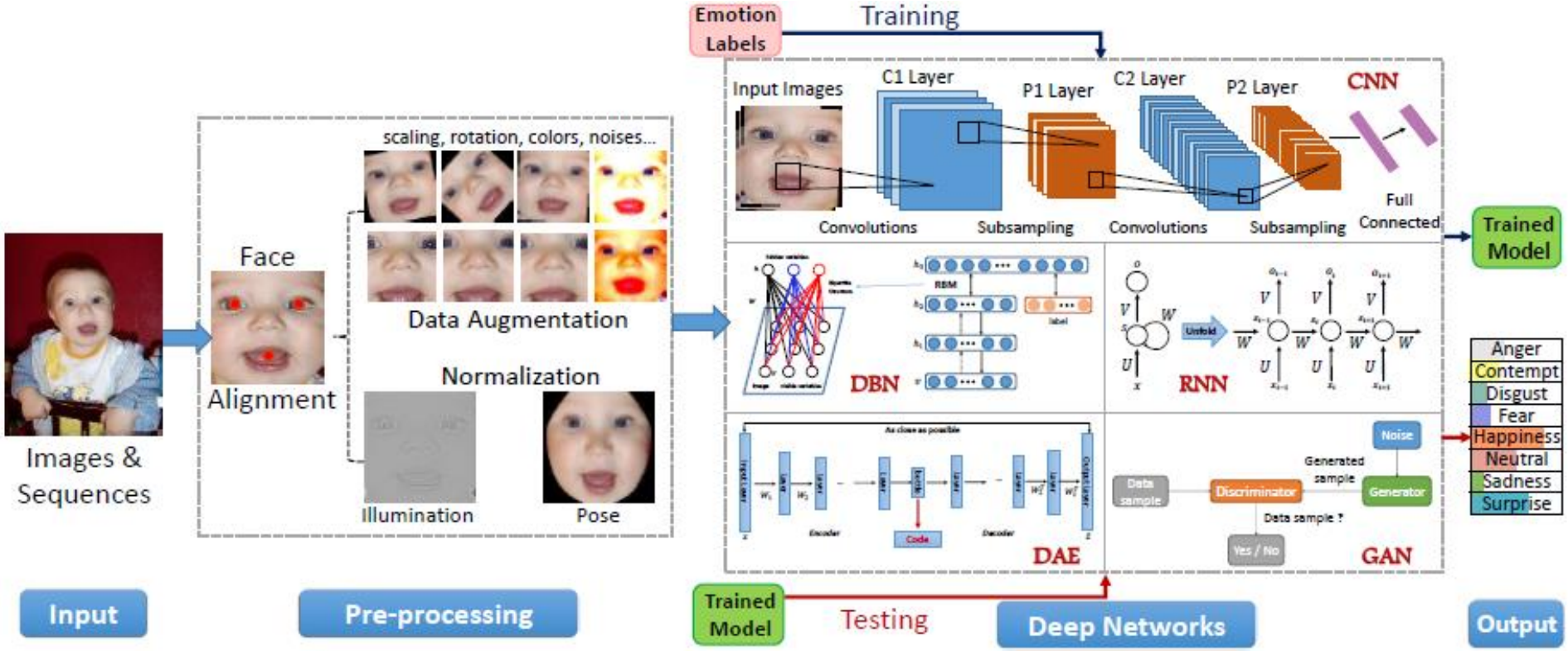
Facial Expression Recognition

Expression Recognition in 2012…



Facial Expression Recognition

Expression Recognition in 2020!



What happened?

Shan Li and Weihong Deng, *Deep Facial Expression Recognition: A Survey*, 2018.

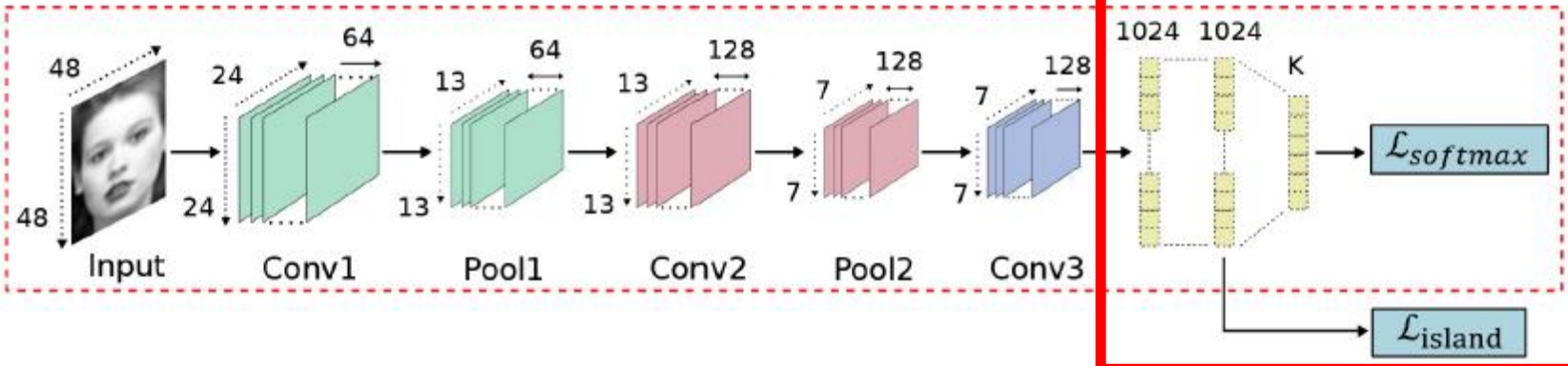
Facial Expression Recognition

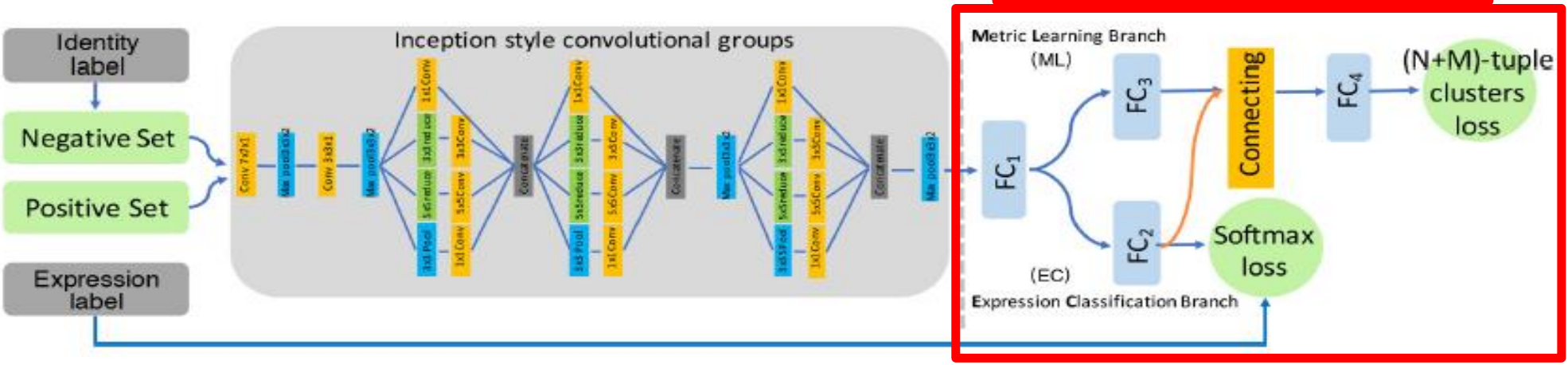
Larger datasets

State-of-the-art models

Complexity of emotions

Only these portions are different!





Underlying feature extraction mechanism remains the same

Shan Li and Weihong Deng, *Deep Facial Expression Recognition: A Survey*, 2018.

With that said, how do we recognize emotions?

Facial Expression Recognition

We will look at two methods-

1. Deep Separable Layers → Xception module
2. Improvement Techniques

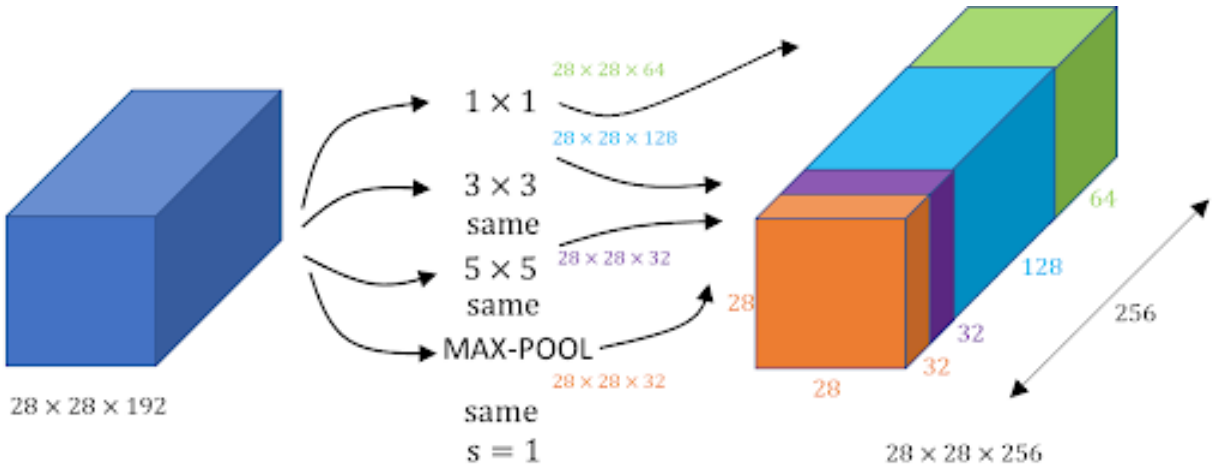
↓

Finetuning
Transfer-Learning
Class Reweighting

Facial Expression Recognition

Xception module

Derived from the famous Inception architecture



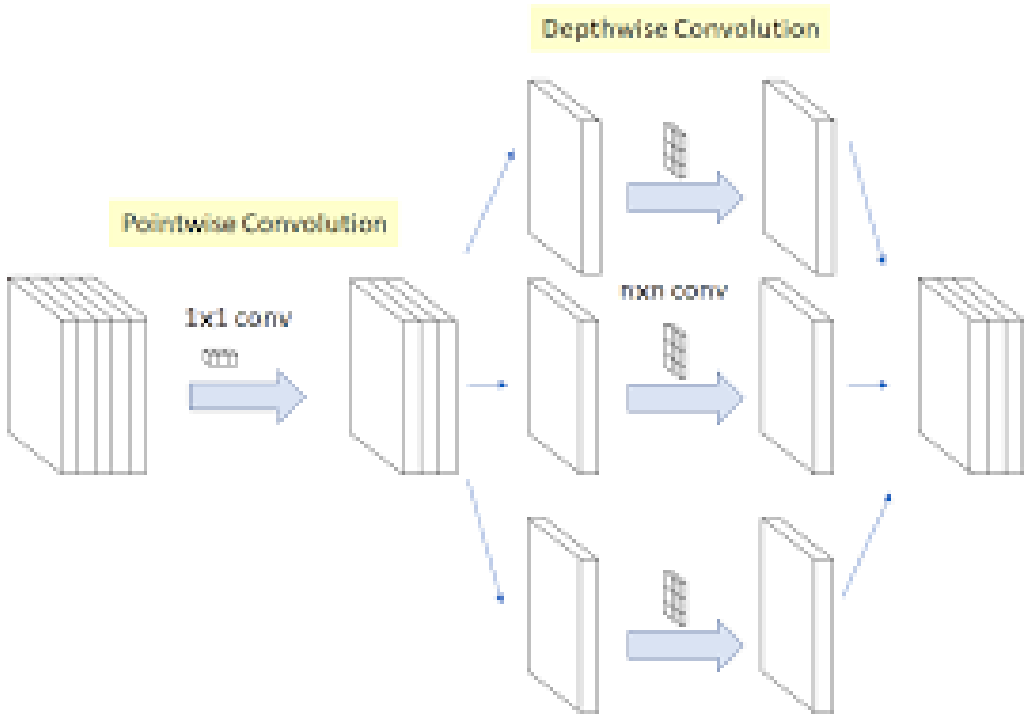
Beneficial for learning richer and dense representations

Split convolutions as per filter sizes

François Chollet, *Xception: Deep Learning with Depthwise Separable Convolutions*, 2017.

Facial Expression Recognition

Xception module



Derived from the famous Inception architecture

Beneficial for learning richer and dense representations

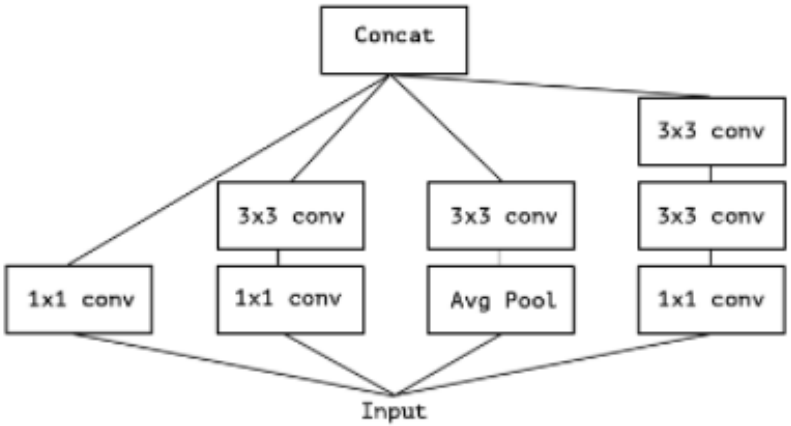
*Split convolutions as per **channels***

François Chollet, *Xception: Deep Learning with Depthwise Separable Convolutions*, 2017.

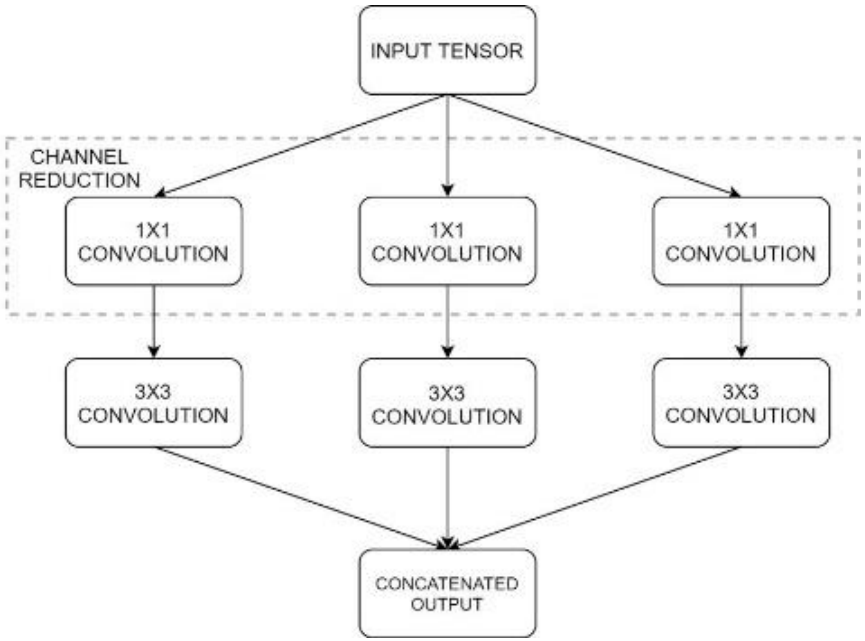
Facial Expression Recognition

Xception module

Inception



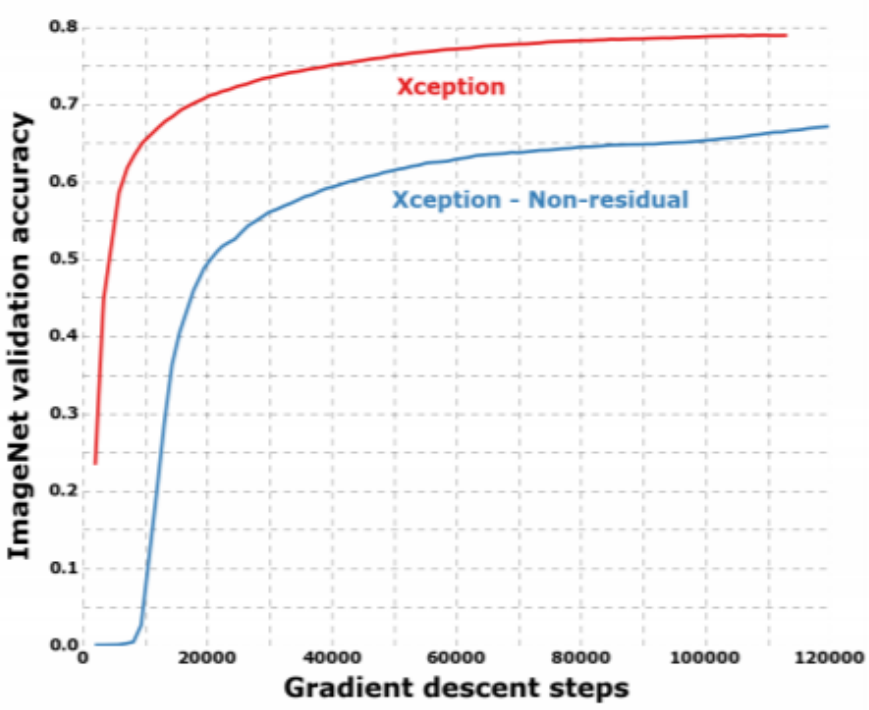
Xception

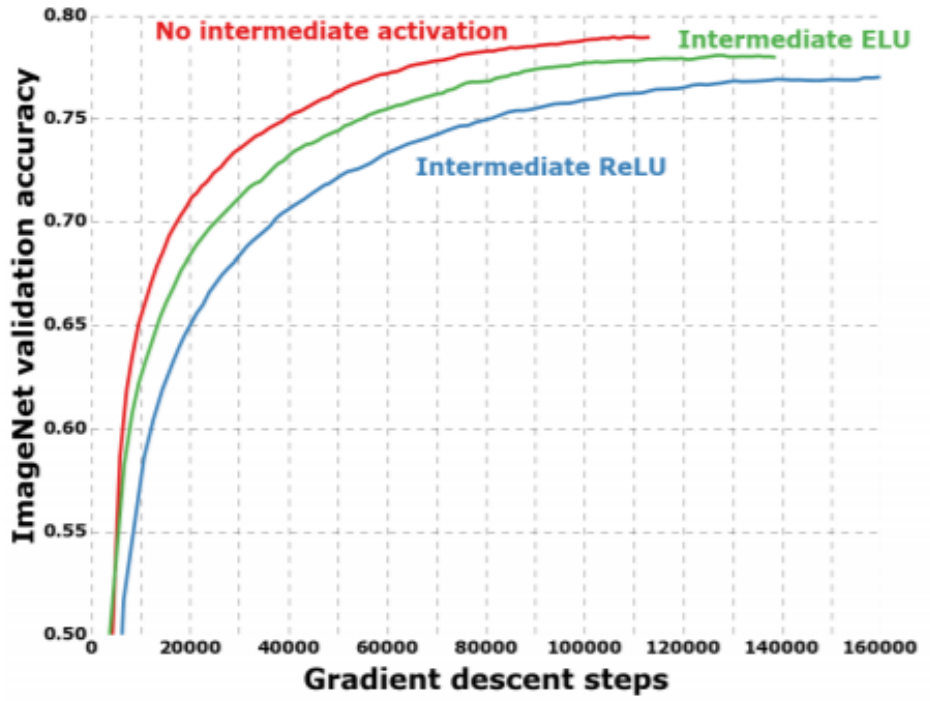


François Chollet, *Xception: Deep Learning with Depthwise Separable Convolutions*, 2017.

Facial Expression Recognition

Xception module





François Chollet, *Xception: Deep Learning with Depthwise Separable Convolutions*, 2017.

Facial Expression Recognition

Xception module

- You will implement Xception for Task-1
- Note that only the layers of the model need to be implemented
- Keep a track of your loss and accuracy metrics
- Visualize as much as you can!

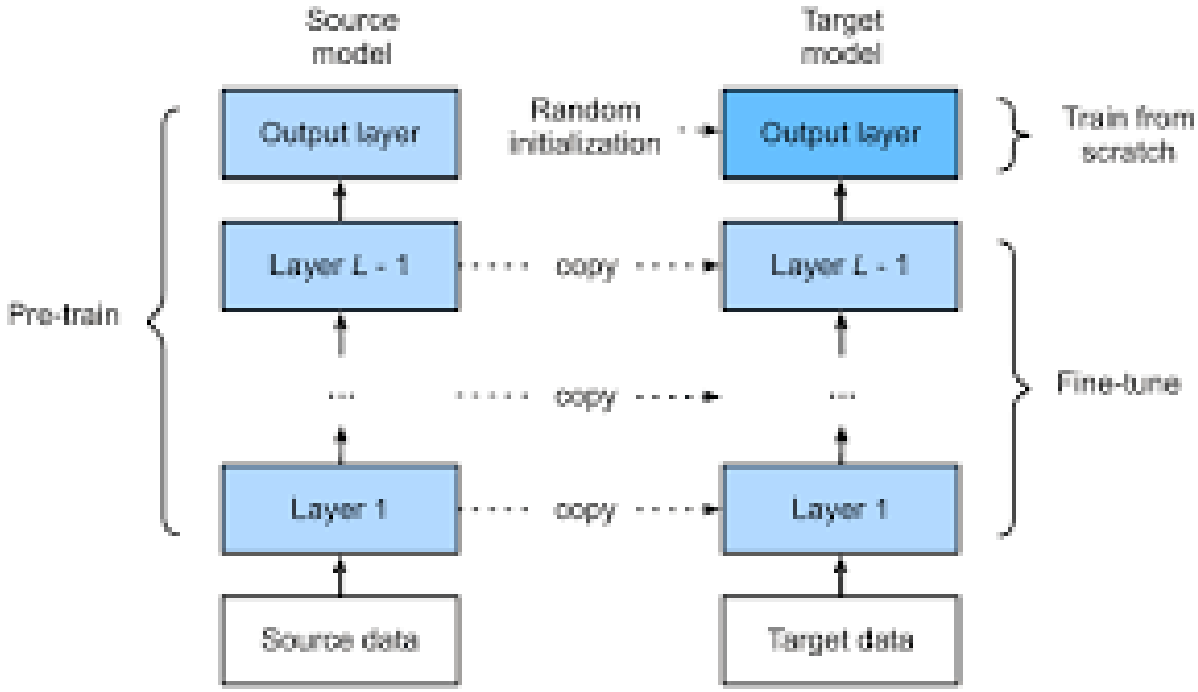
François Chollet, *Xception: Deep Learning with Depthwise Separable Convolutions*, 2017.

Facial Expression Recognition

Improvement Techniques

1. Finetuning

Train only a portion of the model



Avoid training large models from scratch!

Tips-

- Slower learning rate
- Freezing earlier layers
- Train for less epochs
- Monitor performance frequently

Facial Expression Recognition

Improvement Techniques

1. Finetuning

Advantages-

- Works well for very large models
- Scalable to larger datasets

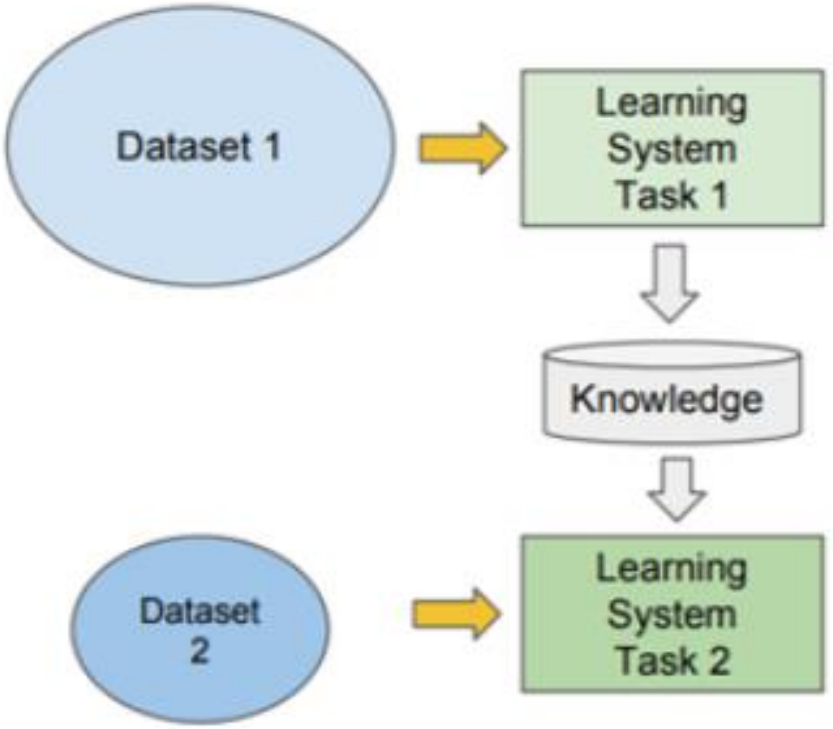
Disadvantages-

- May degrade performance when model not initialized well
- Only applicable to structurally similar datasets

Facial Expression Recognition

Improvement Techniques

2. Transfer-Learning



Similar to finetuning but for different domains

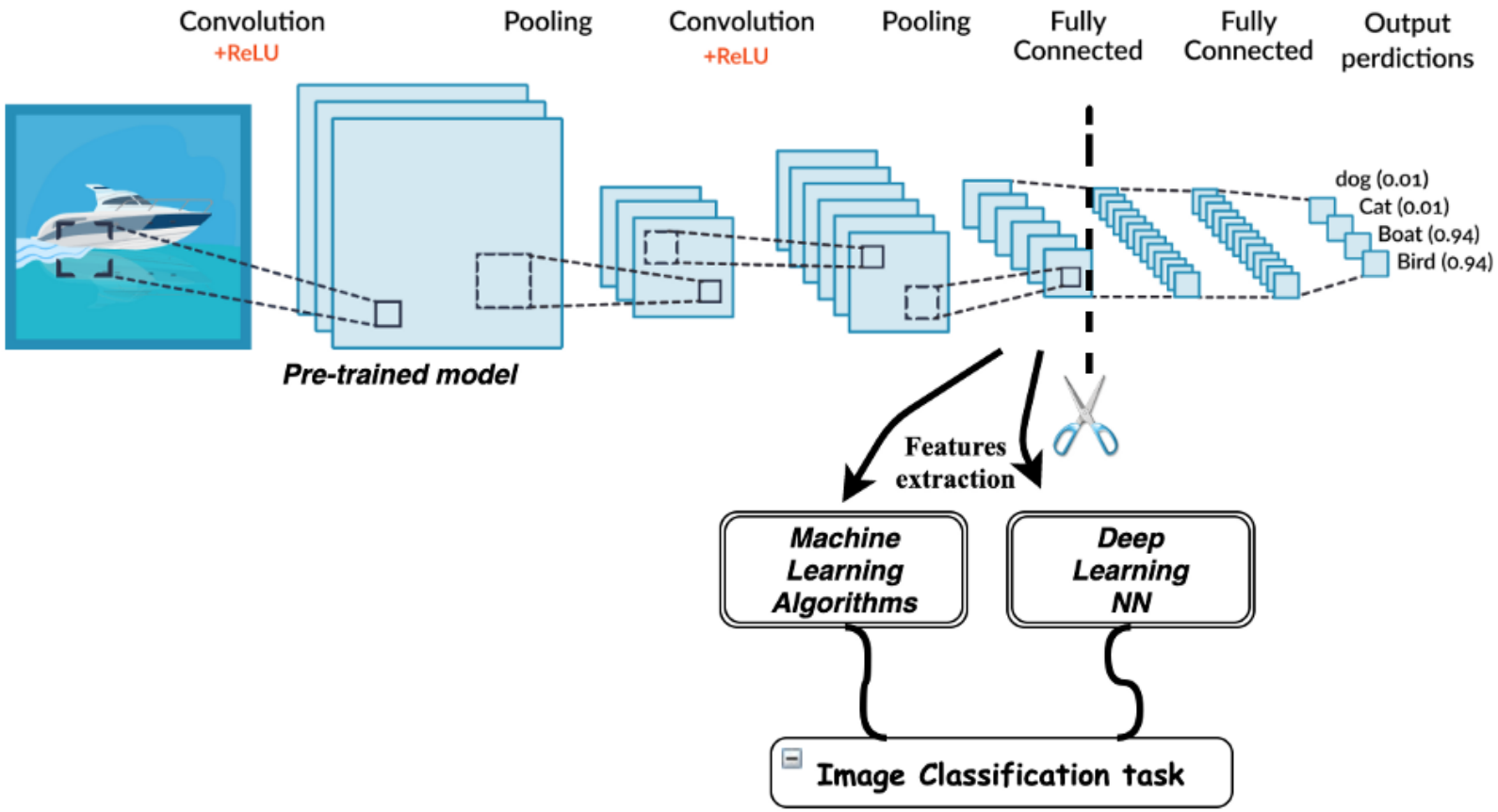
Source task

Target task

Facial Expression Recognition

Improvement Techniques

2. Transfer-Learning



Facial Expression Recognition

Improvement Techniques

2. Transfer-Learning

Advantages-

- Works well for complex datasets
- Knowledge can be transferred to smaller models

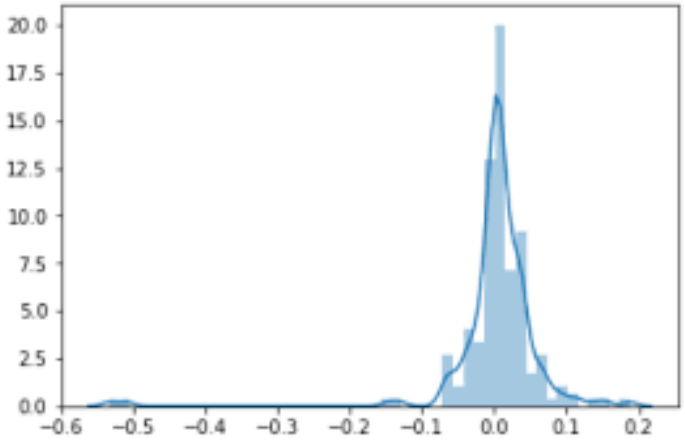
Disadvantages-

- Not suitable for different datasets (dog vs cat)
- Model may suffer from catastrophic forgetting

Facial Expression Recognition

Improvement Techniques

2. Class-Reweighting



Datasets are often biased towards class distributions

Reweight classes as per a criterion-

- Frequency of samples
- Train/Test split
- Application at hand

Just because the model was trained on 10 classes does not mean it will classify all 10 correctly

Improvement Techniques

2. Class-Reweighting

Advantages-

- Addresses dataset bias
- Only trains the model on essential data samples

Disadvantages-

- May not be effective when all classes are extremely large
- Difficult to come up with a practical weighting scheme

Today

- Tutorial (~ 1.5 hours)
  - Part-1: Facial Expression Recognition
  - Part-2: Facial Action Units
  - Part-3: Overview of PyTorch
- Break (~ 5 minutes)
- Open question hour (~ 0.5 hour)
  - Questions, questions, questions…

In case you have a question-

- Raise hand and unmute Mic
- Type in Chat

Facial Action Units

Often difficult to naively interpret emotions

Which of the above two is sad? Why?

Facial Action Units

Expressions can have different semantic meaning

Sad

Bored

Lazy

Tired

Disappointed

A face in general does not convey a lot of information when compared to its individual components

Facial Action Units

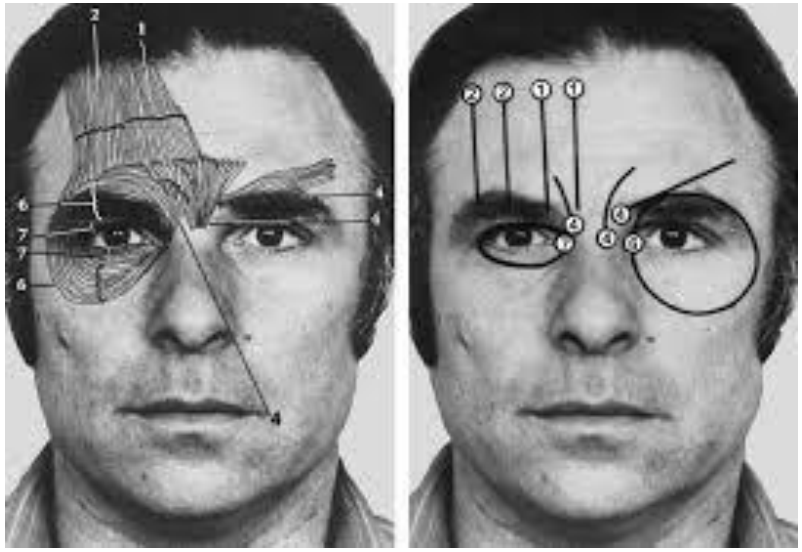
We as humans already know which portions of the face indicate a particular emotion

Sad? Why?

Happy? Why?

Facial Action Units

Ekman's Theory



Certain regions and movements of the face characterize emotions

Paul Ekman. *Facial action coding system (facs)*. A human face, 2002.

Facial Action Units

Ekman's Theory

Group each such unit by its action on the face, hence **Facial Action Unit (FAU)**

Paul Ekman. *Facial action coding system (facs)*. A human face, 2002.

Facial Action Units

Ekman's Theory

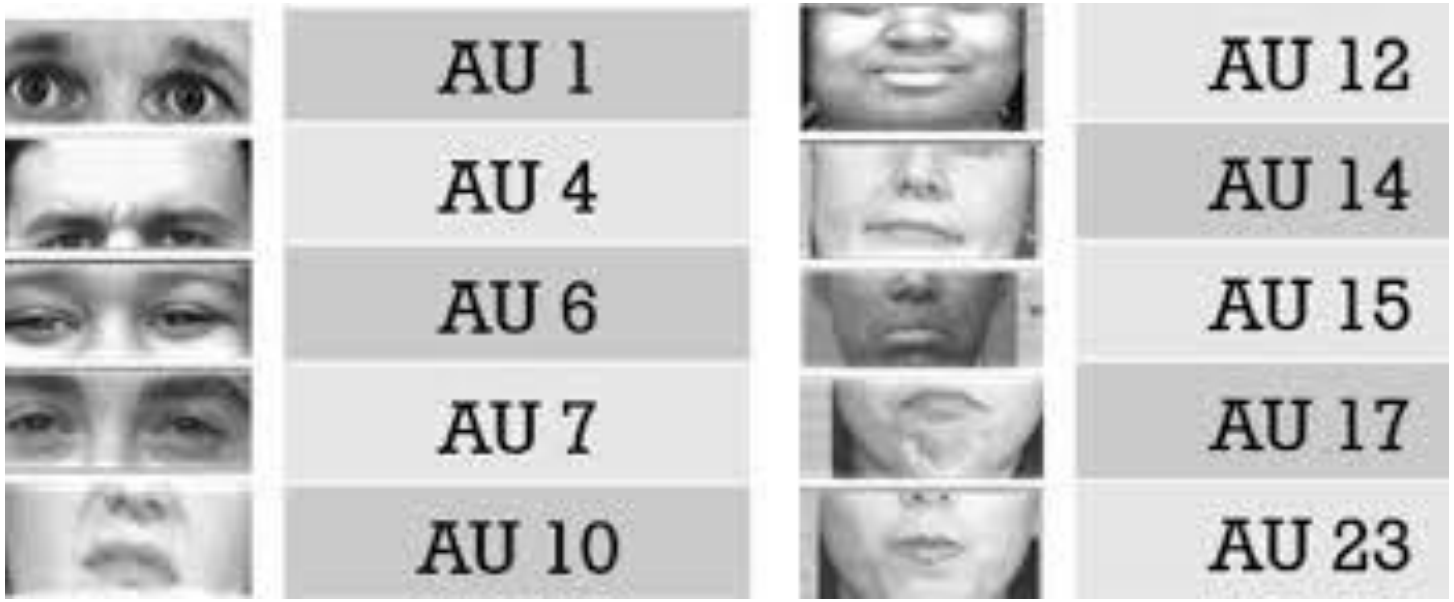
Turns out there are quite a lot of FAUs, do we need to recognize all of them?!

Paul Ekman. *Facial action coding system (facs)*. A human face, 2002.

Facial Action Units

Ekman's Theory

Main FAUs, other emotions are a combination of these units



Not really!

Paul Ekman. *Facial action coding system (facs)*. A human face, 2002.

With that said, how do we recognize FAUs?

Facial Action Units

FAU Recognition

We generally focus on FAU intensity estimation

This helps us quantify the degree of emotion

Only need to look at this region, how cool is that!

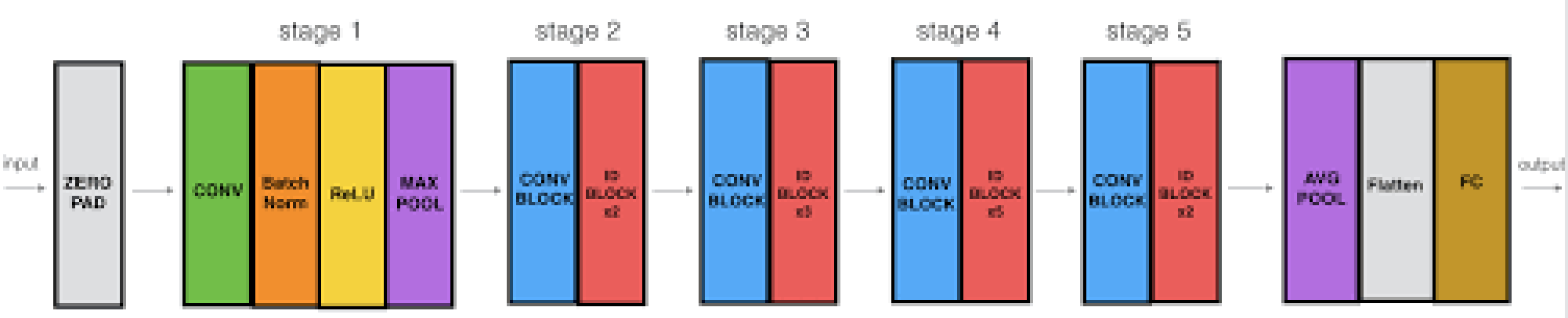
Very Sad! Sad Neutral Happy Very Happy!

Facial Action Units

FAU Recognition

Lets look at a simple and effective method

The ResNet



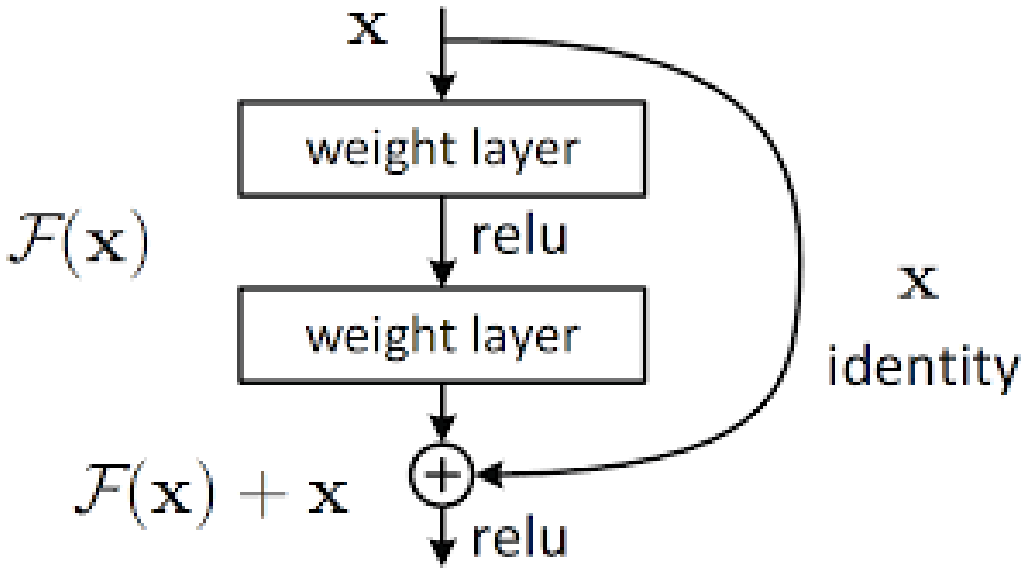
Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. *Deep residual learning for image recognition*. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

Facial Action Units

FAU Recognition

Lets look at a simple and effective method

The ResNet



Residual connections help with-

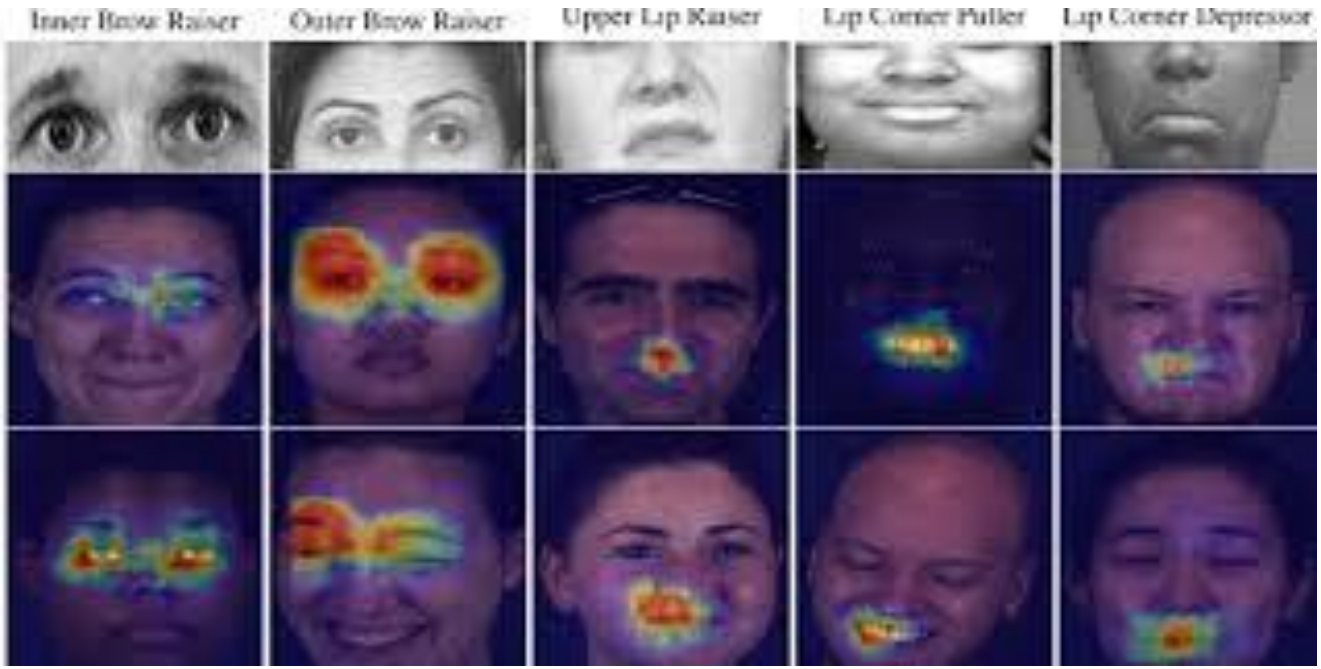
- Richer representations
- Vanishing gradients
- Consistency in local features

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. *Deep residual learning for image recognition*. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

Facial Action Units

FAU Recognition

FAU intensities using deep ResNets



Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. *Deep residual learning for image recognition*. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

Facial Action Units

FAU Recognition

You will be implementing and improving ResNet for Task 2

You need to implement the bottleneck block along with deconvolutional layers

Please read the starter code carefully to avoid errors!

Model is pretrained and will only work if your implementation is correct

Visualize as much as you can!

Facial Action Units

Can we improve FAU recognition?

1. Paper-1: D-pattnet: Dynamic patch-attentive deep network for action unit detection.

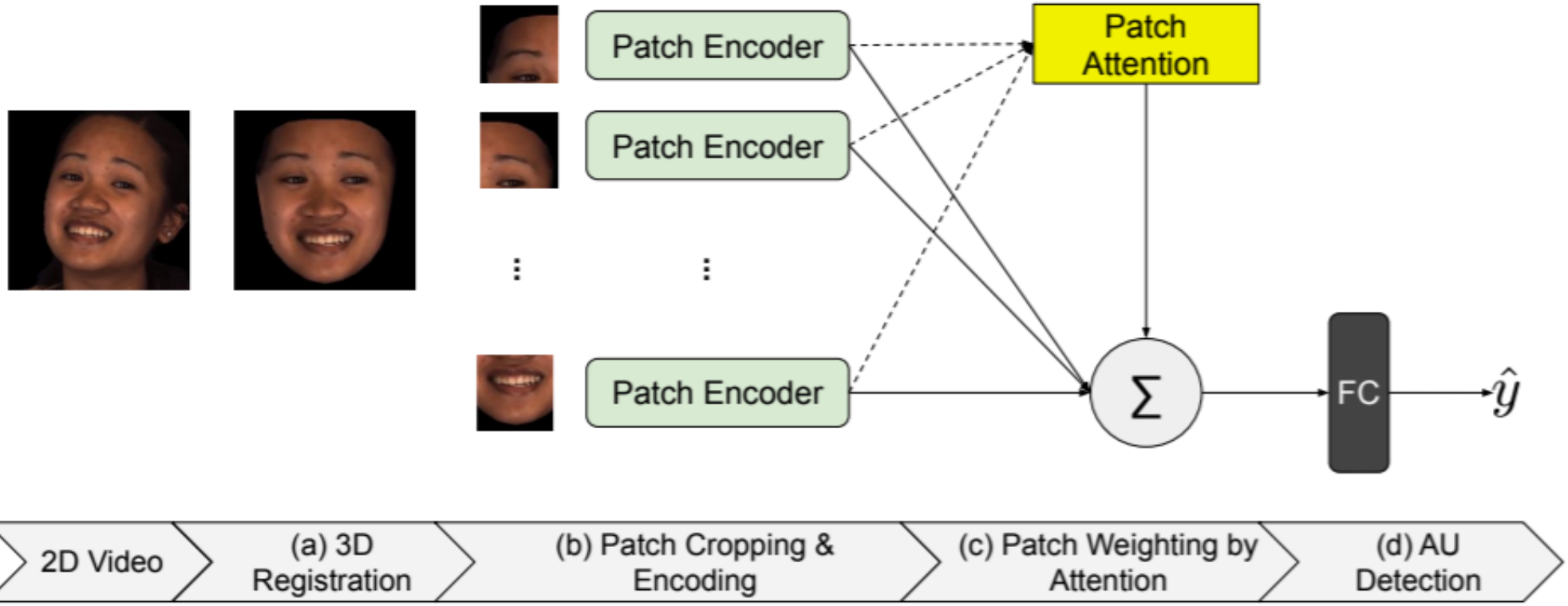
2. Paper-2: Attention-based facial behavior analytics in social communication.

3. Paper-3: Large margin loss for learning facial movements from pseudo-emotions

4. Paper-4: Unmasking the devil in the details: What works for deep facial action coding?

Facial Action Units

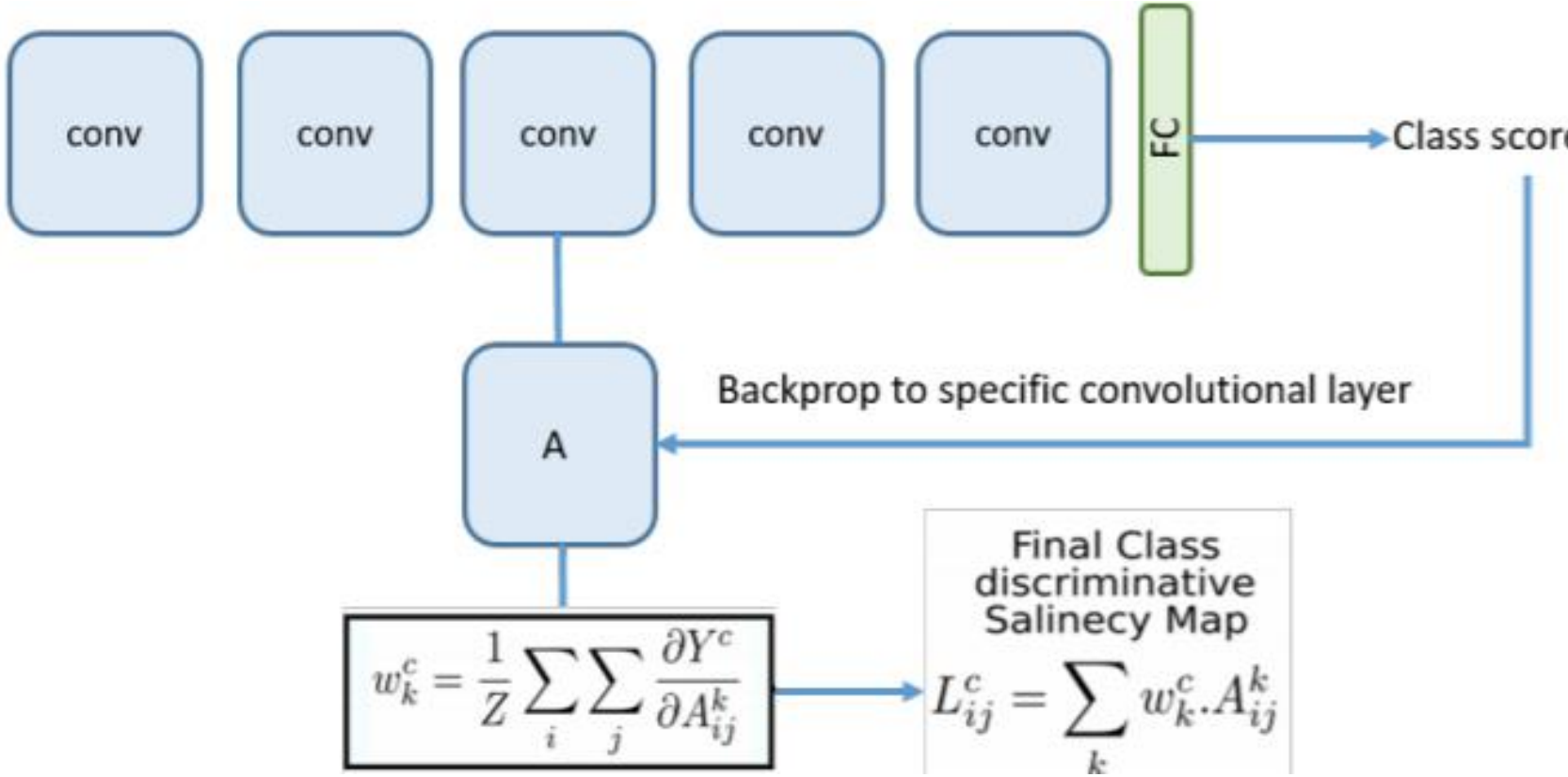
Paper-1



Itir Onal Ertugrul, Le Yang, Laszlo A Jeni, and Jerey F Cohn. D-pattnet: *Dynamic patch-attentive deep network for action unit detection*. Frontiers in computer science, 1:11, 2019.

Facial Action Units

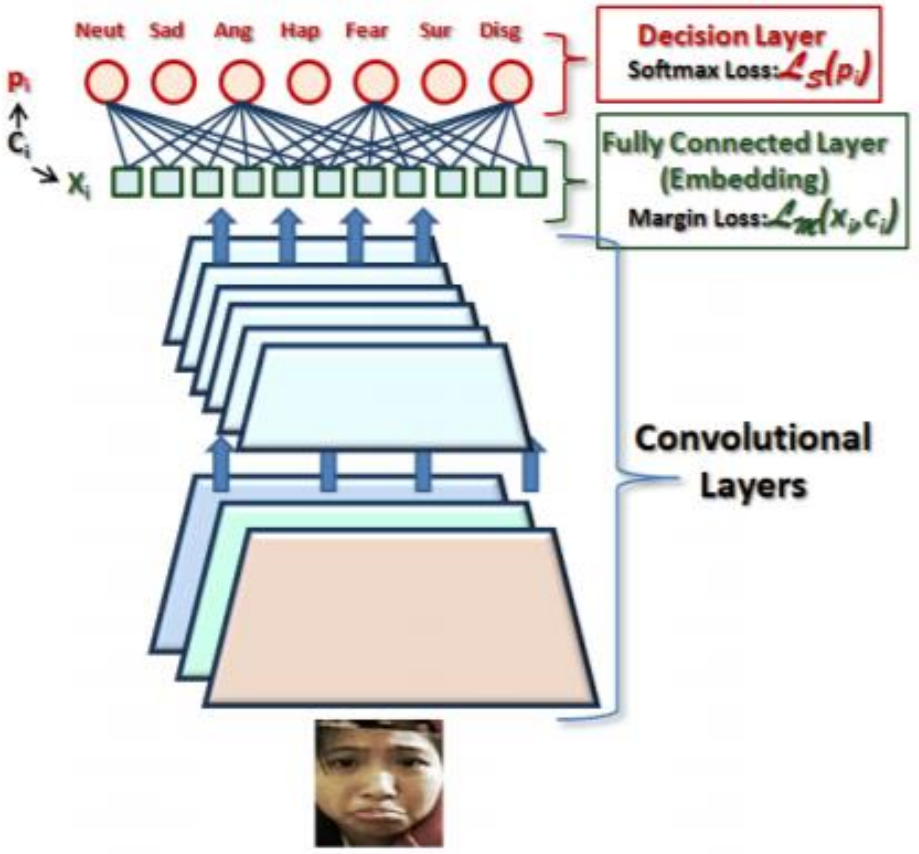
Paper-2



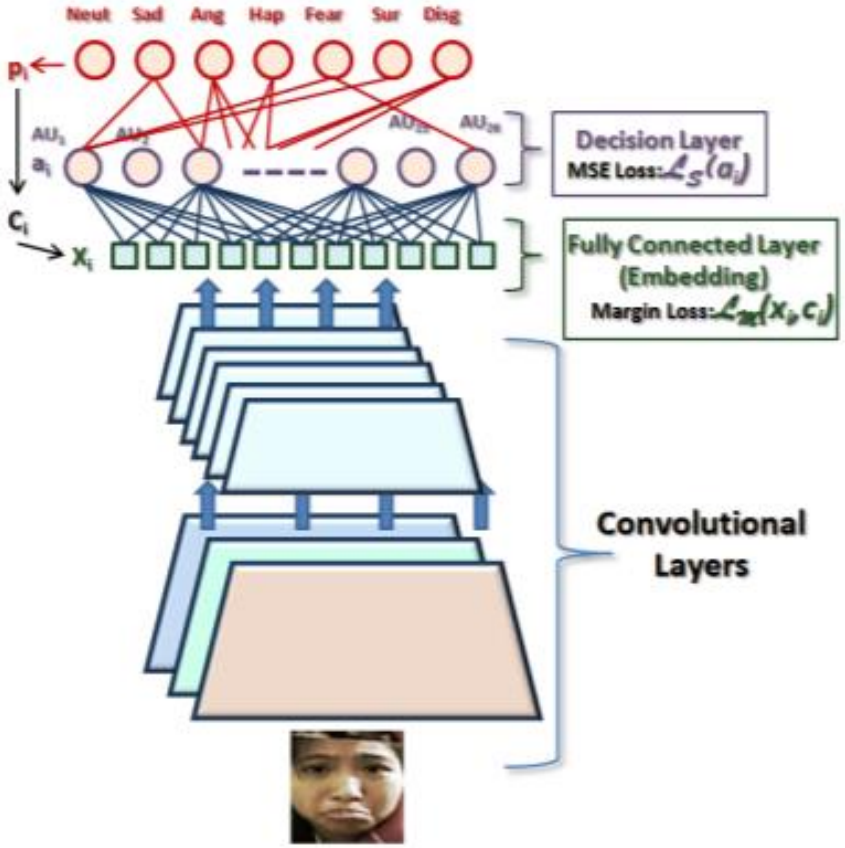
Lezi Wang, Chongyang Bai, Maksim Bolonkin, Judee Burgoon, Norah Dunbar, VS Subrahmanian, and Dimitris N Metaxas. *Attention-based facial behavior analytics in social communication*. In 30th BMVC, 2020.

Facial Action Units

Paper-3



Learning expressions

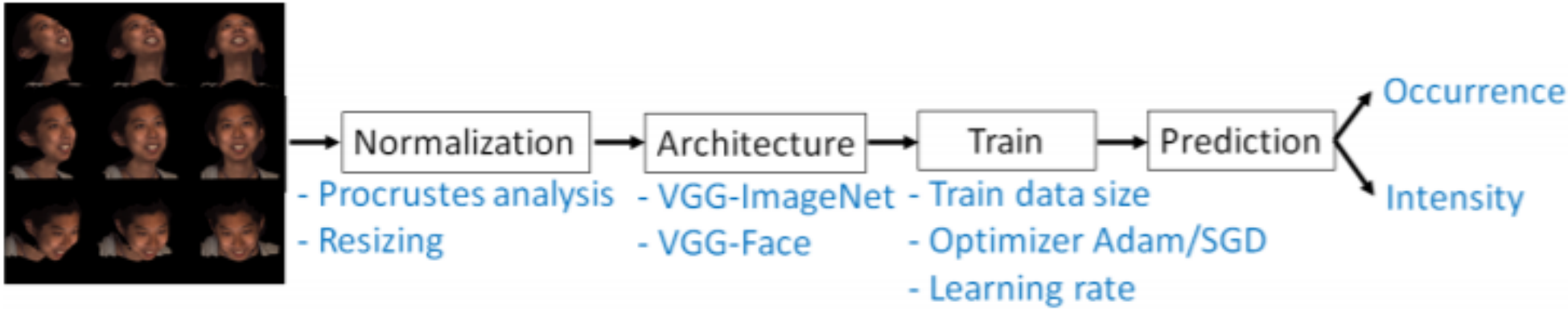


Learning AU from pseudo emotions

Andrei Racoviteanu, Mihai-Sorin Badea, Corneliu Florea, Laura Florea, and Constantin Vertan. *Large margin loss for learning facial movements from pseudo-emotions*.

Facial Action Units

Paper-4



Koichiro Niinuma, Laszlo A. Jeni, Itir• Onal Ertugrul, and Jerey F. Cohn. *Unmasking the devil in the details: What works for deep facial action coding?* In 30th British Machine Vision Conference 2019.

Today

- Tutorial (~ 1.5 hours)
  - Part-1: Facial Expression Recognition
  - Part-2: Facial Action Units
  - Part-3: Overview of PyTorch
- Break (~ 5 minutes)
- Open question hour (~ 0.5 hour)
  - Questions, questions, questions…

In case you have a question-

- Raise hand and unmute Mic
- Type in Chat

Overview of PyTorch

How do we train a model?

$$\theta^* = \arg\min_{\theta} \sum_{(x,y)\in\mathcal{D}} ||f_\theta(x) - y||_2^2$$

$$\theta^* = \arg\min_{\theta} \sum_{(x,y)\in\mathcal{D}} ||f_\theta(x) - y||_2^2$$

gradient descent

neural network

Overview of PyTorch

What is PyTorch?

Python library for-

- Defining neural networks
- Automatically computing gradients

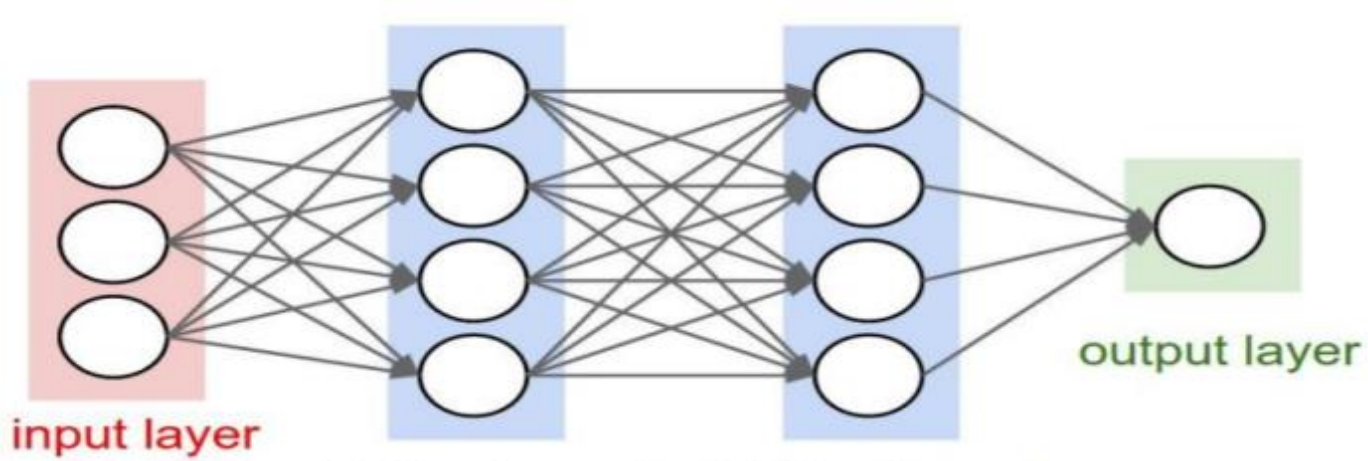
PyTorch alternatives-

- Tensorflow
- JAX
- Chainer
- …

Essentially all do the same thing

Overview of PyTorch

How does PyTorch work?

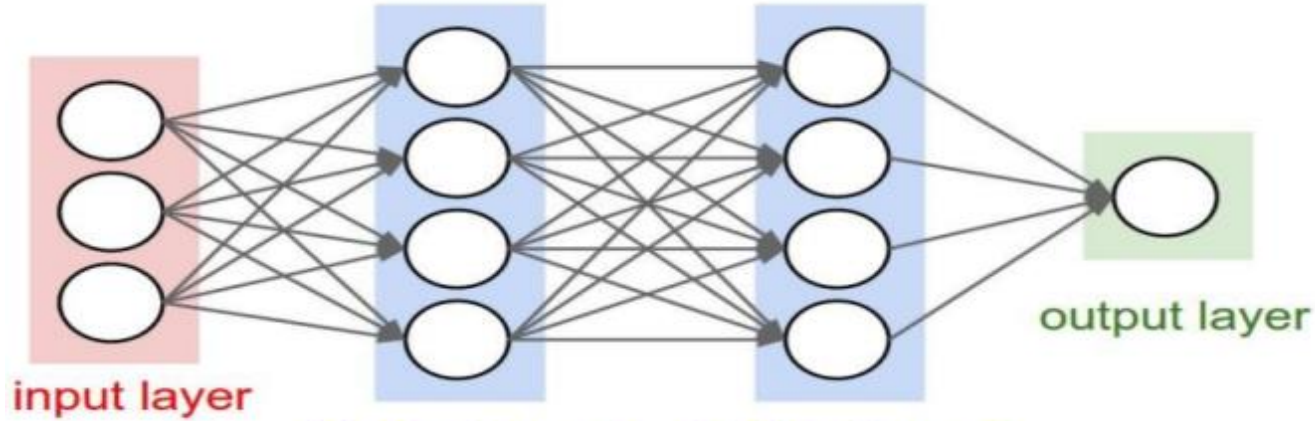


$$h_1 = \sigma(W_1 x) \qquad h_2 = \sigma(W_2 h_1) \qquad y = \sigma(W_3 h_2)$$

Forward Pass

Overview of PyTorch

How does PyTorch work?



You define	$h_1 = \sigma(W_1 x)$	$h_2 = \sigma(W_2 h_1)$	$y = \sigma(W_3 h_2)$
PT computes	$\frac{\partial y}{\partial W_1} = \frac{\partial y}{\partial h_2}\frac{\partial h_2}{\partial h_1}\frac{\partial h_1}{\partial W_1}$	$\frac{\partial y}{\partial W_2} = \frac{\partial y}{\partial h_2}\frac{\partial h_2}{\partial W_2}$	$\frac{\partial y}{\partial W_3}$

[picture from Stanford's CS231n]

Backward Pass

Overview of PyTorch

PyTorch Tutorial in Colab

Link also provided in Project-B handout so that you can go through it at your own pace

https://colab.research.google.com/drive/1r4omUxckDpfhMs_py1yWWPprWy7R8rzj?usp=sharing

Today

- Tutorial (~ 1.5 hours)
  - Part-1: Facial Expression Recognition
  - Part-2: Facial Action Units
  - Part-3: Overview of PyTorch
- Break (~ 5 minutes)
- Open question hour (~ 0.5 hour)
  - Questions, questions, questions…

In case you have a question-

- Raise hand and unmute Mic
- Type in Chat

Today

- Tutorial (~ 1.5 hours)
  - Part-1: Facial Expression Recognition
  - Part-2: Facial Action Units
  - Part-3: Overview of PyTorch
- Break (~ 5 minutes)
- Open question hour (~ 0.5 hour)
  - Questions, questions, questions…

In case you have a question-

- Raise hand and unmute Mic
- Type in Chat

Open Question Hour

Questions..

In case you have a question-

- Raise hand and unmute Mic
- Type in Chat