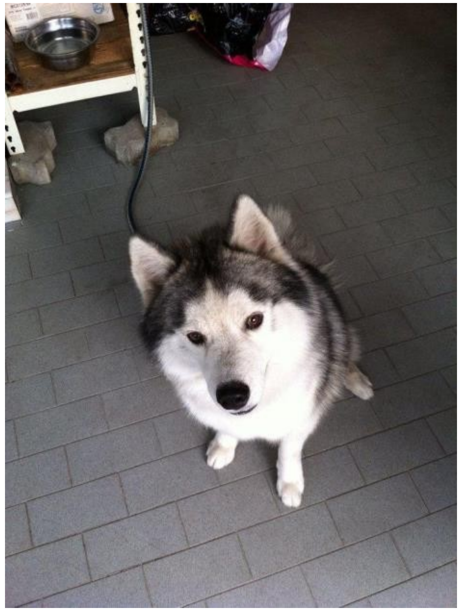
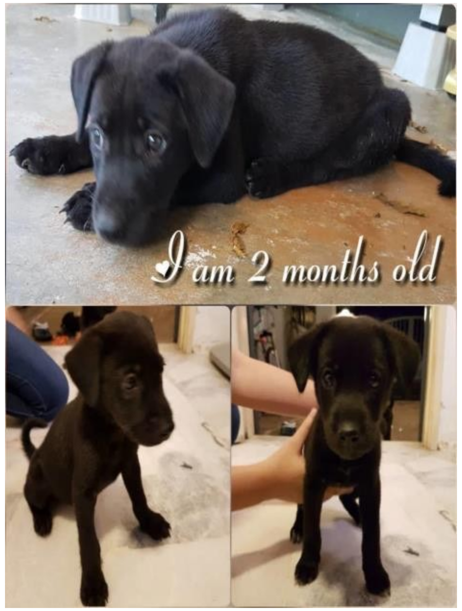
Topic: Pawpularity Contest

https://www.kaggle.com/competitions/petfinder-pawpularity-score

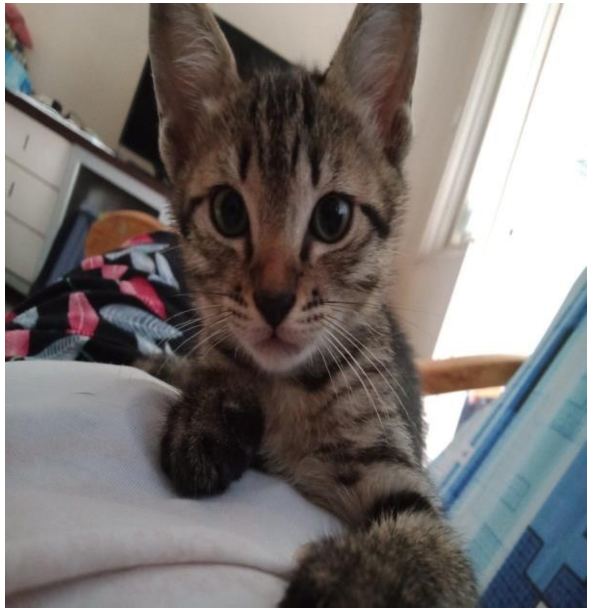
Pawpularity Contest (Dogs & Cats)



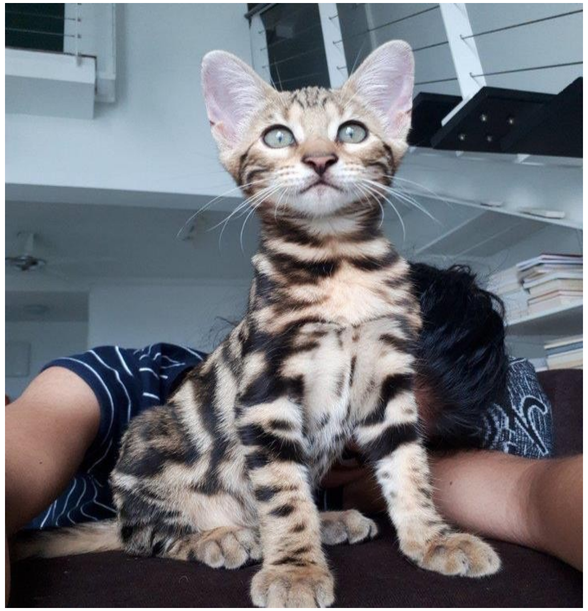
Value: 100



Value: 77



Value: 45



Value: 100

- Predict the pawpularity of pet photos.
- An image regression task to predict values from 0 to 100.

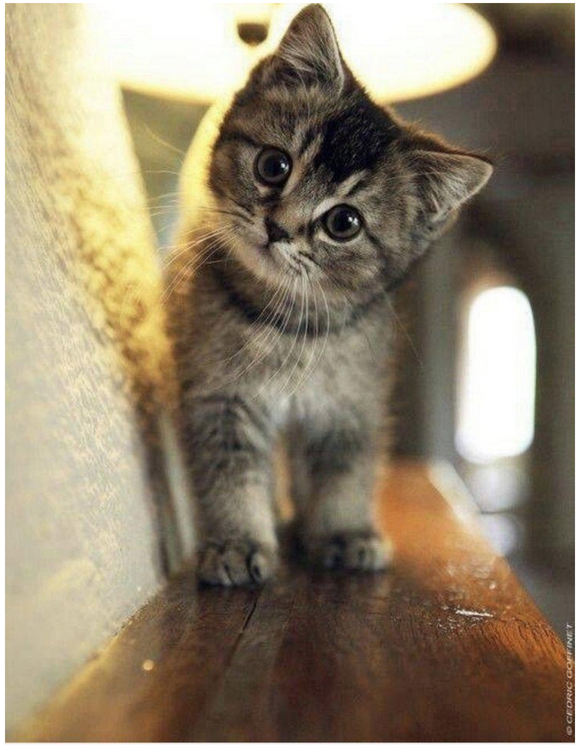
Training data

- Contain 9912 images with pawpularity scores and metadata (auxiliary labels).

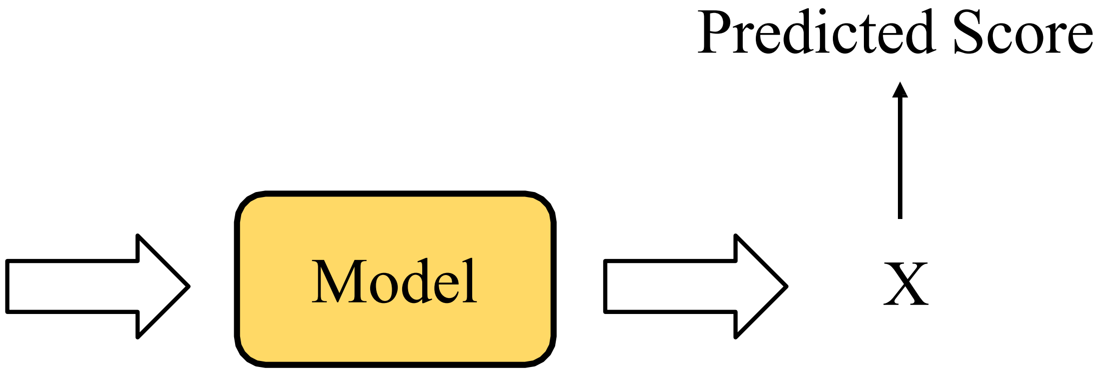
- Pawpularity scores contain scores from 0 to 100:

- Metadata contains 12 manual labels like the below:
  - Focus: Pets are stands out against uncluttered background.
  - Eyes: Both eyes are facing front or near-front.
  - Face: Clear Face.
  - Near: Single pet taking up significant portion of photo.
  - Action: Pet in the middle of an action.
  - ……

Data Introduction



Input



Label

Score: 63

Focus: 1

Eyes: 1

Face: 1

Near: 1

Action: 0

⋮ : 0

Testing Data

- Contain 6800 images without any photo labels

- Evaluation: Root Mean Square Error (RMSE)

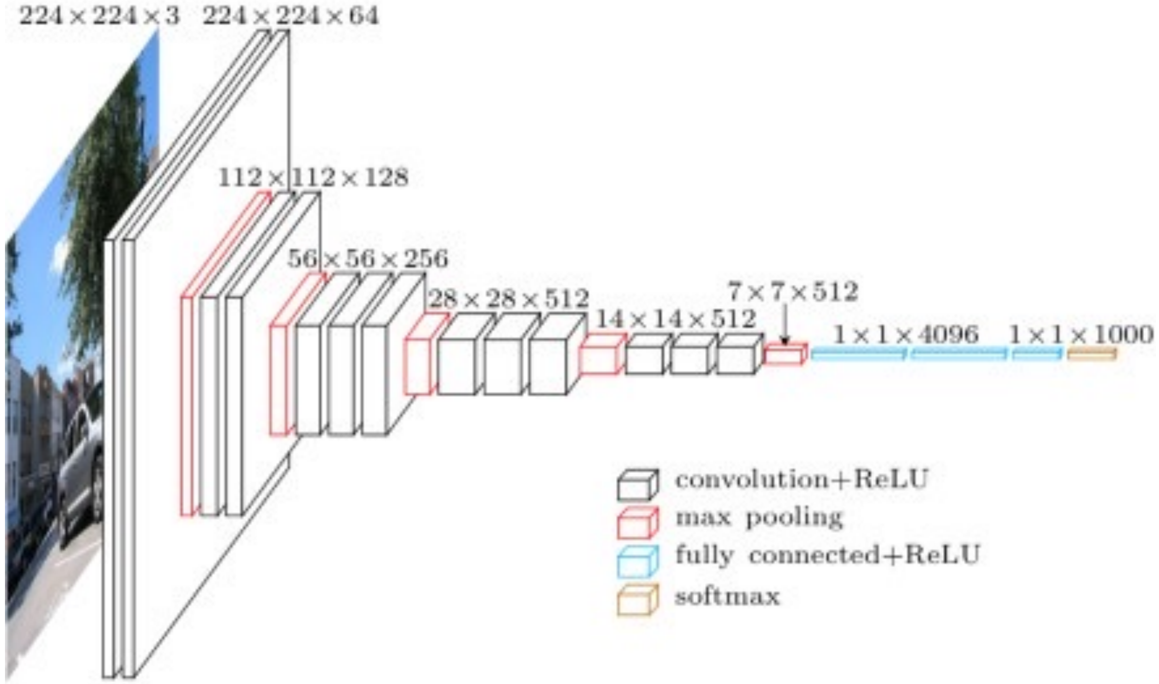
$$RMSE = \sqrt{\frac{\sum_{n=1}^{N}(\hat{y}_n - y_n)^2}{N}}$$

$\hat{y}_n$: Predictions

y_n: Label

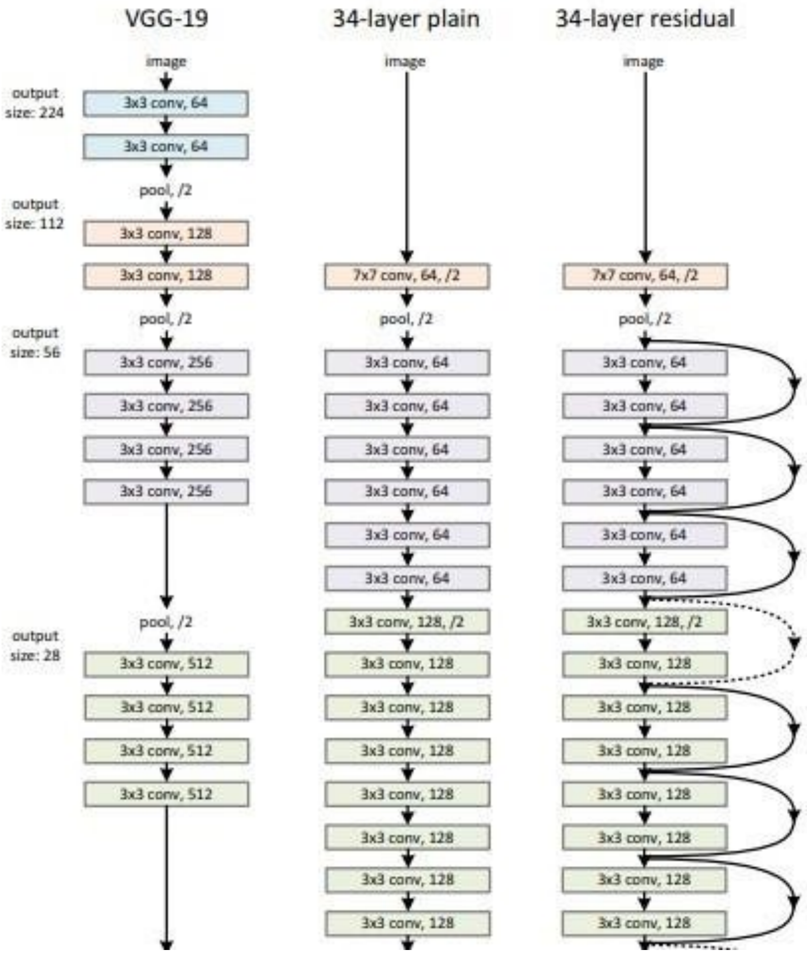
- You only need to predict pawpularity scores

Baseline Model (1/4)



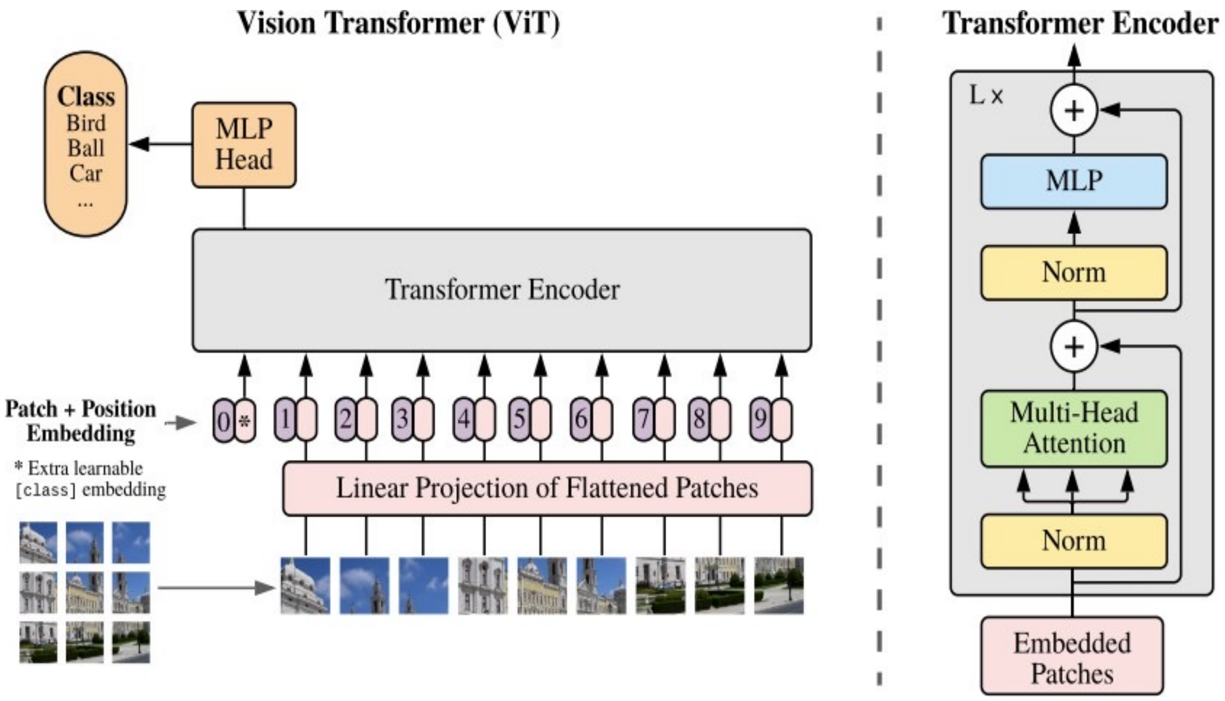
- VGG16 and VGG19
- A basic model for image classification
- Pre-trained model is available

Baseline Model (2/4)



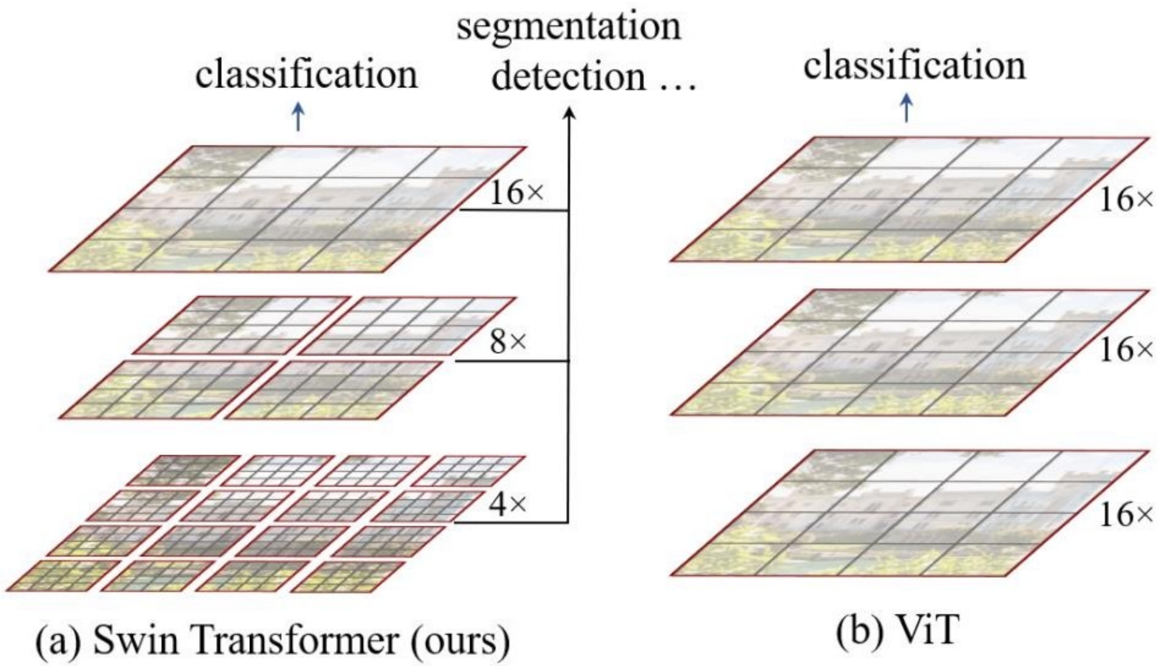
- ResNet34, ResNet50, ResNet101, ResNet152
- A basic model for image classification
- Pre-trained model is available

Baseline Model (3/4)



- Vision Transformer (ViT)
- A transformer-based method for image classification
- It might lead to out of memory

Baseline Model (4/4)



- Swin Transformer
- A transformer-based method for image classification
- A token-efficient transformer

How to Make a Fair Comparison

- It is quite important to show your novelty in this project
- Please make a fair comparison to show your contribution

Model	Error Rate (%)	Param (M)
ResNet-34	21.81	21.5
ResNet-50	20.74	23.9
ResNet-101	19.87	42.8
ResNet-50 + **X**	19.87	30

- **X** is your proposed module
- You can show that your method has its contribution
- You can also compare runtime and FLOPs