

Student 1: Wu Yidu **EIDs:** yiduwu2

Student 2: Guo Yuqi **EIDs:** yuqiguo4

Kaggle Competition: PetFinder.my - Pawpularity Contest

Kaggle Team Name: CS5489-Group23

CS5489 - Course Project (2021A)

Due date: See canvas site.

Possible Projects

For the course project, you may select **one** of the following competitions on Kaggle **or** define your own course project:

[**PetFinder.my - Pawpularity Contest: Predict the popularity of shelter pet photos**](#)

A picture is worth a thousand words. But did you know a picture can save a thousand lives? Millions of stray animals suffer on the streets or are euthanized in shelters every day around the world. You might expect pets with attractive photos to generate more interest and be adopted faster. But what makes a good picture? With the help of data science, you may be able to accurately determine a pet photo's appeal and even suggest improvements to give these rescue animals a higher chance of loving homes.

PetFinder.my is Malaysia's leading animal welfare platform, featuring over 180,000 animals with 54,000 happily adopted. PetFinder collaborates closely with animal lovers, media, corporations, and global organizations to improve animal welfare.

Currently, PetFinder.my uses a basic Cuteness Meter to rank pet photos. It analyzes picture composition and other factors compared to the performance of thousands of pet profiles. While this basic tool is helpful, it's still in an experimental stage and the algorithm could be improved.

In this competition, you'll analyze raw images and metadata to predict the "Pawpularity" of pet photos. You'll train and test your model on PetFinder.my's thousands of pet profiles. Winning versions will offer accurate recommendations that will improve animal welfare.

If successful, your solution will be adapted into AI tools that will guide shelters and rescuers around the world to improve the appeal of their pet profiles, automatically enhancing photo quality and recommending composition improvements. As a result, stray dogs and cats can find their "furever" homes much faster. With a little assistance from the Kaggle community, many precious lives could be saved and more happy families created.

Top participants may be invited to collaborate on implementing their solutions and creatively improve global animal welfare with their AI skills.

G-Research Crypto Forecasting: Use your ML expertise to predict real crypto market data

Over \$40 billion worth of cryptocurrencies are traded every day. They are among the most popular assets for speculation and investment, yet have proven wildly volatile. Fast-fluctuating prices have made millionaires of a lucky few, and delivered crushing losses to others. Could some of these price movements have been predicted in advance?

In this competition, you'll use your machine learning expertise to forecast short term returns in 14 popular cryptocurrencies. We have amassed a dataset of millions of rows of high-frequency market data dating back to 2018 which you can use to build your model. Once the submission deadline has passed, your final score will be calculated over the following 3 months using live crypto data as it is collected.

The simultaneous activity of thousands of traders ensures that most signals will be transitory, persistent alpha will be exceptionally difficult to find, and the danger of overfitting will be considerable. In addition, since 2018, interest in the cryptomarket has exploded, so the volatility and correlation structure in our data are likely to be highly non-stationary. The successful contestant will pay careful attention to these considerations, and in the process gain valuable insight into the art and science of financial forecasting.

G-Research is Europe's leading quantitative finance research firm. We have long explored the extent of market prediction possibilities, making use of machine learning, big data, and some of the most advanced technology available. Specializing in data science and AI education for workforces, Cambridge Spark is partnering with G-Research for this competition.

Student-defined Course Project

The goal of the student-defined project is to get some hands-on experience using the course material on your own research problems. Keep in mind that there will only be about 4 weeks to do the project, so the scope should not be too large. Following the major themes of the course, here are some general topics for the project:

- *regression* (supervised learning) - use regression methods (e.g. ridge regression, Gaussian processes) to model data or predict from data.
- *classification* (supervised learning) - use classification methods (e.g., SVM, BDR, Logistic Regression, NNs) to learn to distinguish between multiple classes given a feature vector.
- *clustering* (unsupervised learning) - use clustering methods (e.g., K-means, EM, Mean-Shift) to discover the natural groups in data.
- *visualization* (unsupervised learning) - use dimensionality reduction methods (e.g., PCA, kernel-PCA, non-linear embedding) to visualize the structure of high-dimensional data.

You can pick any one of these topics and apply them to your own problem/data.

- *Can my project be my recently submitted or soon-to-be submitted paper?* If you plan to just turn in the results from your paper, then the answer is no. The project cannot be based on work that you have already done. However, your course project can be based on extending your work. For example, you can try some models introduced in the course on your data/problem.

Before actually doing the project, you need to write a **project proposal** so that we can make sure the project is doable within the 3-4 weeks. I can also give you some pointers to relevant methods, if necessary.

- The project proposal should be at most one page with the following contents: 1) an introduction that briefly states the problem; 2) a precise description of what you plan to do - e.g., What types of features do you plan to use? What algorithms do you plan to use? What dataset will you use? How will you evaluate your results? How do you define a good outcome for the project?
- The goal of the proposal is to work out, in your head, what your project will be. Once the proposal is done, it is just a matter of implementation!
- *You need to submit the project proposal to Canvas 1 week after the Course project is released.*

Groups

Group projects should contain 2 students. To sign up for a group, go to Canvas and under "People", join one of the existing "Project Groups". *For group projects, the project report must state the percentage contribution from each project member.*

Methodology

You are free to choose the methodology to solve the task. In machine learning, it is important to use domain knowledge to help solve the problem. Hence, instead of blindly applying the algorithms to the data you need to think about how to represent the data in a way that makes sense for the algorithm to solve the task.

Kaggle: Kaggle Notebooks

The Kaggle competitions have Kaggle Notebooks enabled, which provide free GPU/TPU computing resources (up to a limit). You can develop your model in the Kaggle Notebook, CS5489 JupyterHub, or on your own computers.

Kaggle: Evaluation on Kaggle

For Kaggle projects, the final evaluation will be performed on Kaggle. Note that for these competitions you need to submit your code via the Kaggle Notebook, which will then generate the submission file for processing.

Project Presentation

Each project group needs to give a presentation at the end of the semester. You will record your presentation and upload it to FlipGrid. The presentation is limited to 5 minutes. You *must* give a presentation. See the details in the "Project Presentations" Canvas assignment.

What to hand in

You need to turn in the following things.

The following files should be uploaded to "Course Project" on Canvas:

1. This ipynb file `CourseProject-2021A.ipynb` with your source code and documentation. **You should write about all the various attempts that you make to find a good solution.** You may also submit .py files, but your documentation should be in the ipynb file.
2. A PDF version of your ipynb file.
3. Presentation slides.
4. (Kaggle projects) Your final submission file to Kaggle.
5. (Kaggle projects) A downloaded copy of your Kaggle Notebook that is submitted to Kaggle. This file should contain the code that generates the final submission file on Kaggle. This code will be used to verify that your Kaggle submission is reproducible.

Other things that need to be turned in:

- Upload your Project presentation to FlipGrid and the submit the URL to the "Project Presentations" assignment on Canvas. See the detailed instructions in the assignment.
- Enter the percentage contribution for each project member using the "Project Group Contribution" assignment on Canvas.
- (Student-defined projects) submit your project proposal to the "Project Proposal" assignment on Canvas. The project proposal is due 1 week after the course project is released. Kaggle projects do not need to submit a proposal.

Grading

The marks of the assignment are distributed as follows:

- 40% - Results using various feature representations, dimensionality reduction methods, classifiers, etc.
- 25% - Trying out feature representations (e.g. adding additional features, combining features from different sources) or methods not used in the tutorials.
- 15% - Quality of the written report. More points for insightful observations and analysis.
- 15% - Project presentation
- 5% - (Kaggle projects) Final ranking on the Kaggle test data, or (student-defined projects) Project proposal.

Late Penalty: 25 marks will be subtracted for each day late.

Group contribution: marks for a group member with less than equal contribution will be deducted according to the following formula:

- Let A% and B% be the percentage contributions for group members Alice and Bob.
A%+B%=100%
 - Let x be the group project marks.
 - If A>B, then Bob's marks will be reduced to be: $x \cdot B/A$
-

YOUR METHODS HERE

```
In [ ]: %matplotlib inline
import IPython.core.display
IPython.core.display.set_matplotlib_formats("svg")
from sklearn import *
import glob
import csv
import string
import pandas as pd
import xgboost as xgb
from scipy import stats
```

```
In [4]: import os
import sys
import math
import re
import gc
from tqdm import tqdm
import pickle
import subprocess
import pandas as pd
import cv2
import matplotlib.pyplot as plt
import numpy as np
from sklearn import metrics
from sklearn import datasets
from sklearn import model_selection
import albumentations
import torch
import torch.nn as nn
import timm
import tez
from tez.callbacks import EarlyStopping
import cudf, cuml
```

```
In [25]: filepath = 'petfinder-pawpularity-score/'
train_df = pd.read_csv(filepath + 'train.csv')
train_df
```

Out[25]:

		Id	Subject Focus	Eyes	Face	Near	Action	Accessory	Gi
0	0007de18844b0dbbb5e1f607da0606e0			0	1	1	1	0	0
1	0009c66b9439883ba2750fb825e1d7db			0	1	1	0	0	0
2	0013fd999caf9a3efe1352ca1b0d937e			0	1	1	1	0	0

		Id	Subject Focus	Eyes	Face	Near	Action	Accessory	Gi
3		0018df346ac9c1d8413cfcc888ca8246		0	1	1	1	0	0
4		001dc955e10590d3ca4673f034feef2		0	0	0	1	0	0
...	
9907		ffbf0383c34dc513c95560d6e1fdb57		0	0	0	1	0	0
9908		ffcc8532d76436fc79e50eb2e5238e45		0	1	1	1	0	0
9909		ffdf2e8673a1da6fb80342fa3b119a20		0	1	1	1	0	0
9910		fff19e2ce11718548fa1c5d039a5192a		0	1	1	1	0	0
9911		fff8e47c766799c9e12f3cb3d66ad228		0	1	1	1	0	0

9912 rows × 14 columns

show some pictures

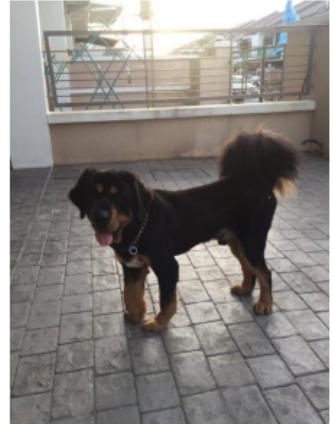
In [6]:

```
import random;
import matplotlib.image as mpimg
rows, cols = 3, 3;
fig, axs = plt.subplots(rows, cols, figsize=(15,15));
fig.subplots_adjust(top = 0.99, bottom=0.01, hspace=0.2, wspace=0.4);
for i in range(rows):
    for j in range(cols):
        random_image = random.randint(0,len(dftrainshow)-1);
        img = mpimg.imread(csv_path+'train/'+dftrainshow['Id'][random_image]+'.jpg');
        axs[i,j].imshow(img);
        axs[i,j].axis('off');
        axs[i,j].set_title(f'Pawpularity: {dftrainshow["Pawpularity"][random_image]}');
```

Pawpularity: 29



Pawpularity: 6



Pawpularity: 35



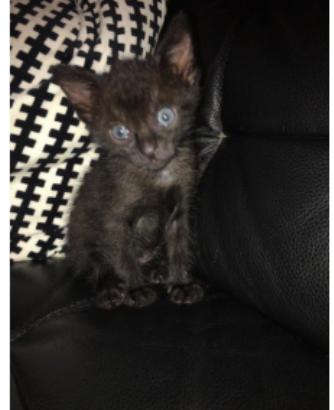
Pawpularity: 42



Pawpularity: 84



Pawpularity: 18



Pawpularity: 48



Pawpularity: 36



Pawpularity: 30



1. Data visualization

Firstly we present the 'Pawpularity' which is the prediction target on the figure.

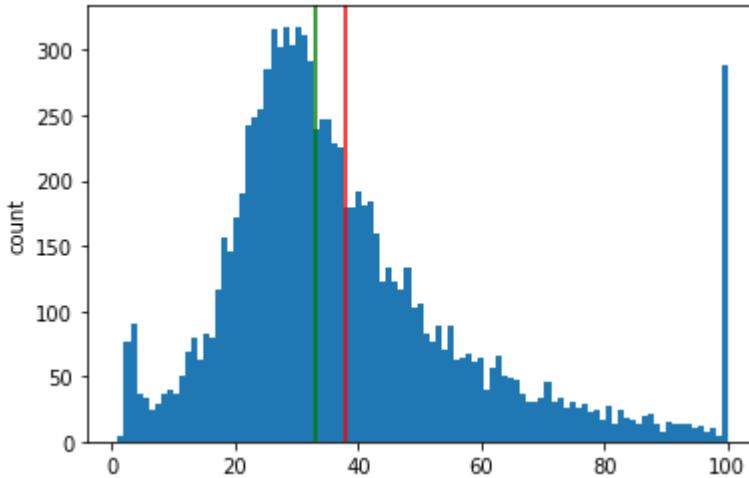
We can observe that the data approximately follows the normal distribution except the smallest ones(which is close to 0) and largest ones(which is close to 100). A conjecture is that the Pawpularity data is generated from people's click-through rate and stay time on each pet image. We can see that some 'very Pawpularity' pictures get a high score(which is the upper limitation), and a low-score clustering which might be the new uploading pictures.

In [8]:

```
import matplotlib.pyplot as plt  
  
plt.hist(train_df['Pawpularity'], bins=100)
```

```
plt.axvline(train_df['Pawpularity'].mean(), color='red')
plt.axvline(train_df['Pawpularity'].median(), color='green')
plt.ylabel('count')
```

Out[8]: Text(0, 0.5, 'count')



Besides the image data, the train.csv also provides several columns of metadata and based on which we draw the violinplot of each column.

And following is the observations:

- Almost data of every column follows normal distribution.
- We didn't find a decisive column which could determine the final Pawpularity score.
- Subject focus: 0>>1, only a few pets stands out against uncluttered background.
- Eyes: 1>0, most pets show their eyes on the picture
- Faces: 1>>0, most pets show their faces on the picture
- Near: 1>>0, most pets takes up more than 50% of photo
- Action: 0>>1, only a few pets have actions
- Accessory: 0>>1, only a few pets accompany physical or digital accessory / prop
- Group: 0>1, most of pictures only have 1 pet
- Collage: 0>>1, only a few pictures are digitally-retouched
- Human: 0>1, no humans in most of pictures
- Occlusion: 0>>1, only a few pictures have specific undesirable objects
- Info: 0>>1, only a few pictures have descriptive texts
- Blur: 0>>1, only a few pictures have 'out of focus or noisy' problem

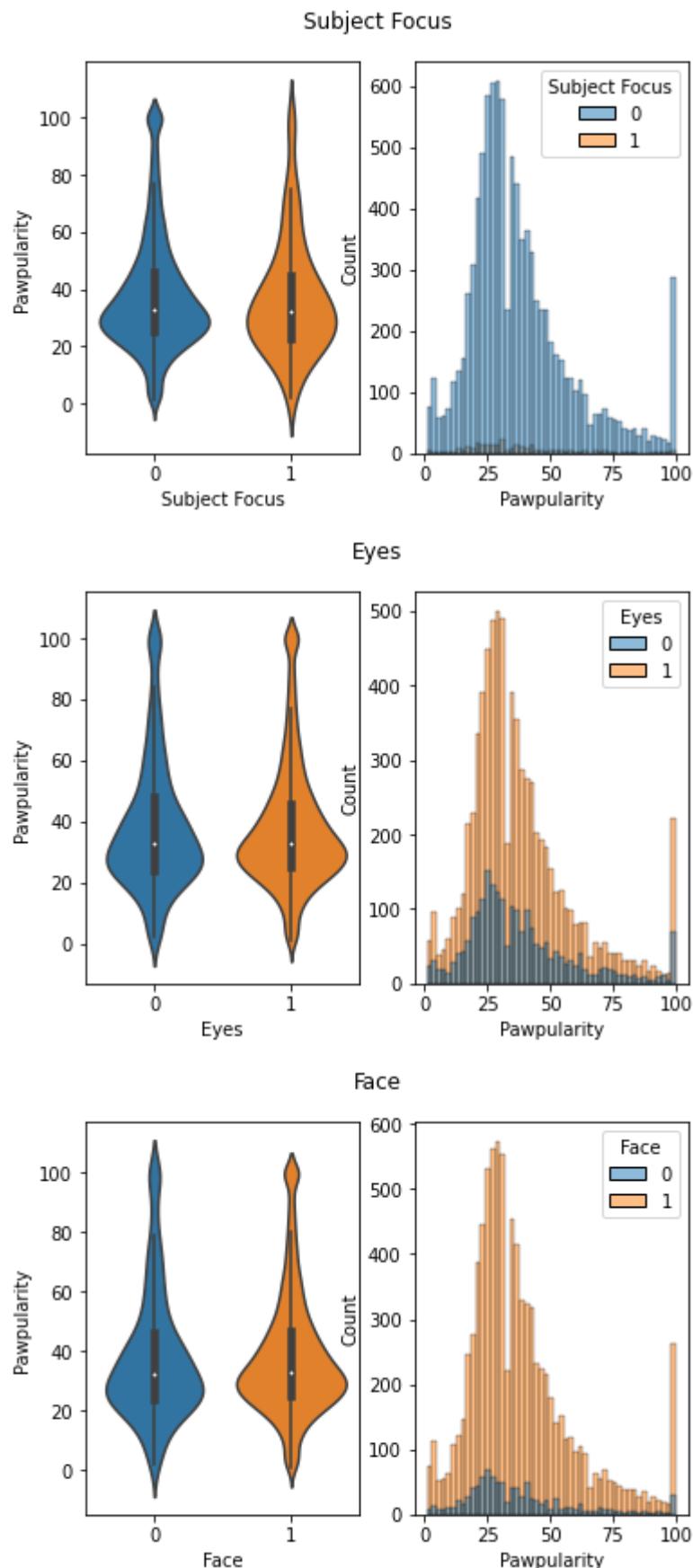
In [9]:

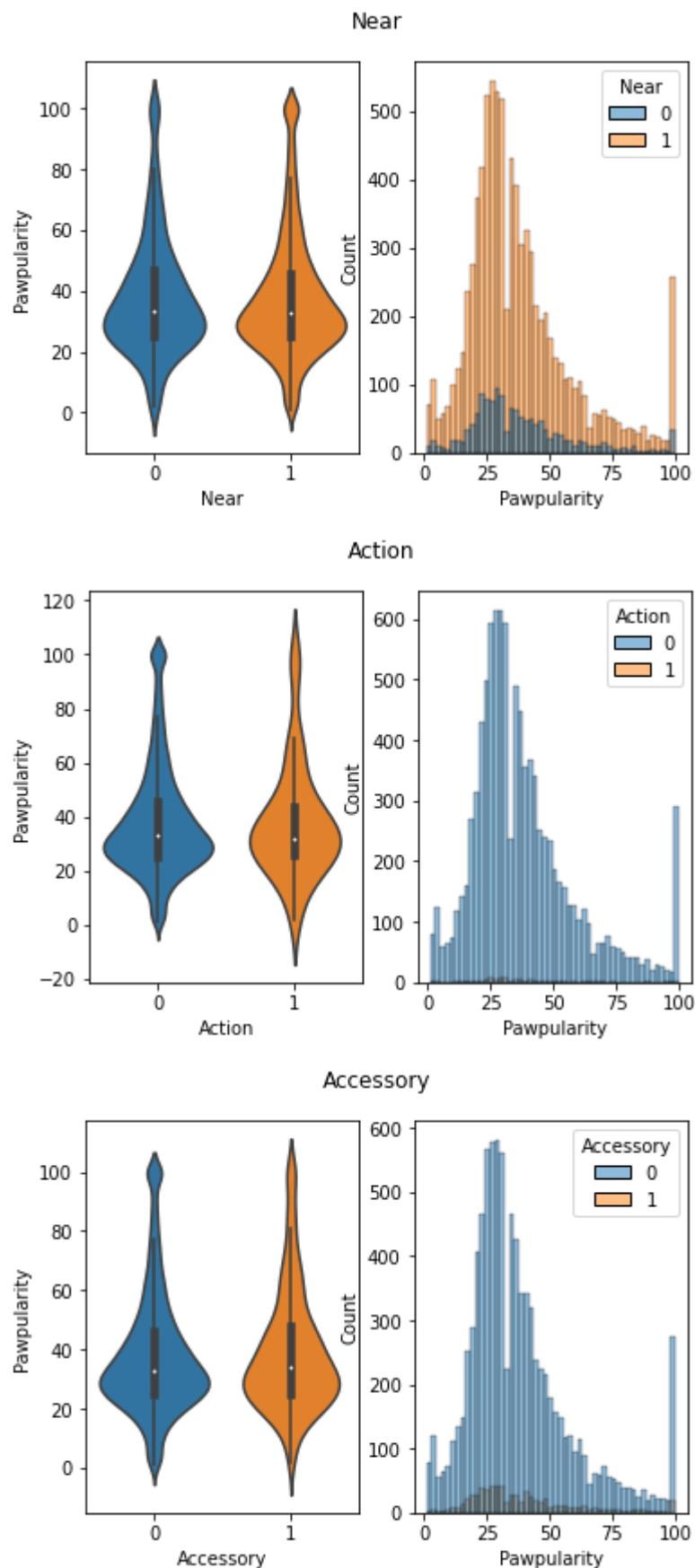
```
import seaborn as sns
cols = ['Subject Focus', 'Eyes', 'Face', 'Near', 'Action', 'Accessory', 'Group'

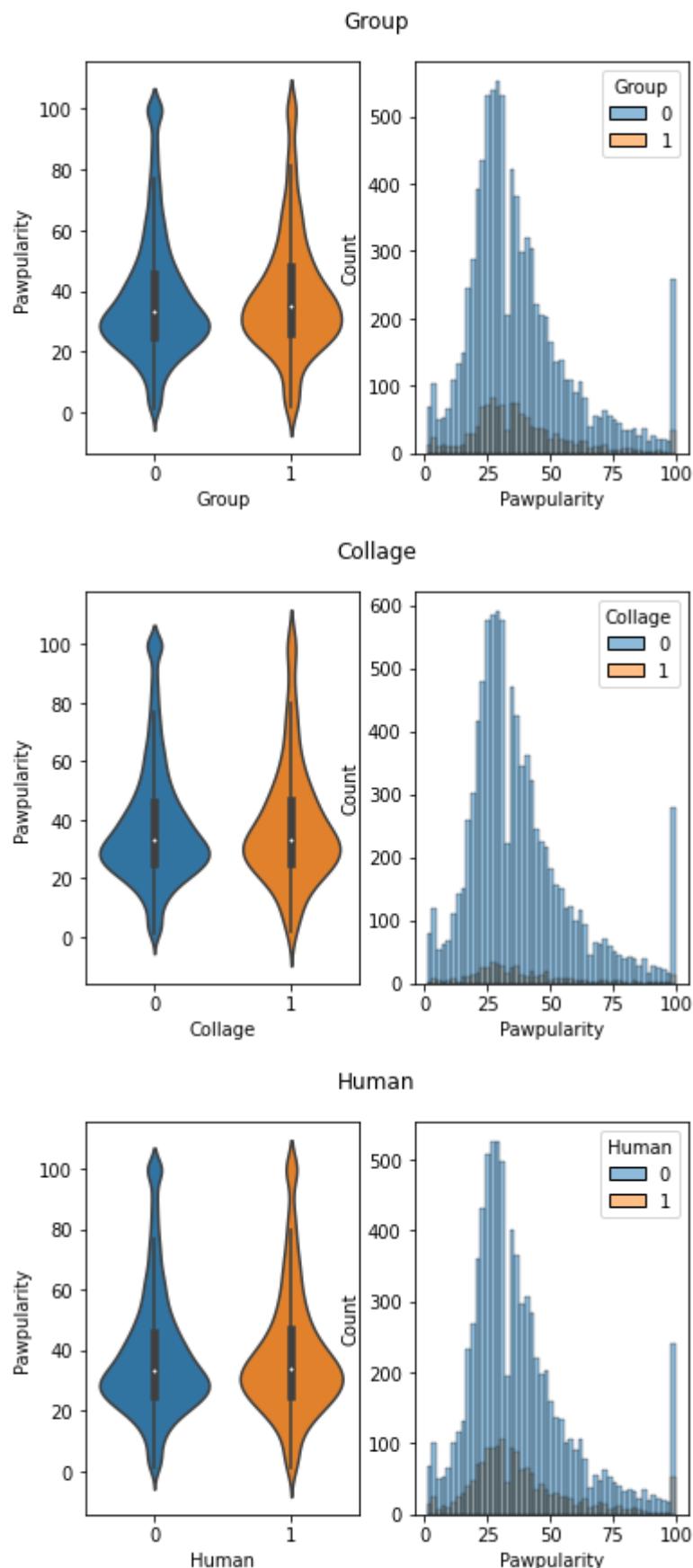
for col in cols:
    fig, ax = plt.subplots(1, 2)
    pos = [0, 1]
    sns.violinplot(x = train_df[col], y = train_df['Pawpularity'], ax=ax[0])
    sns.histplot(data = train_df, hue=col, x ='Pawpularity',ax=ax[1])

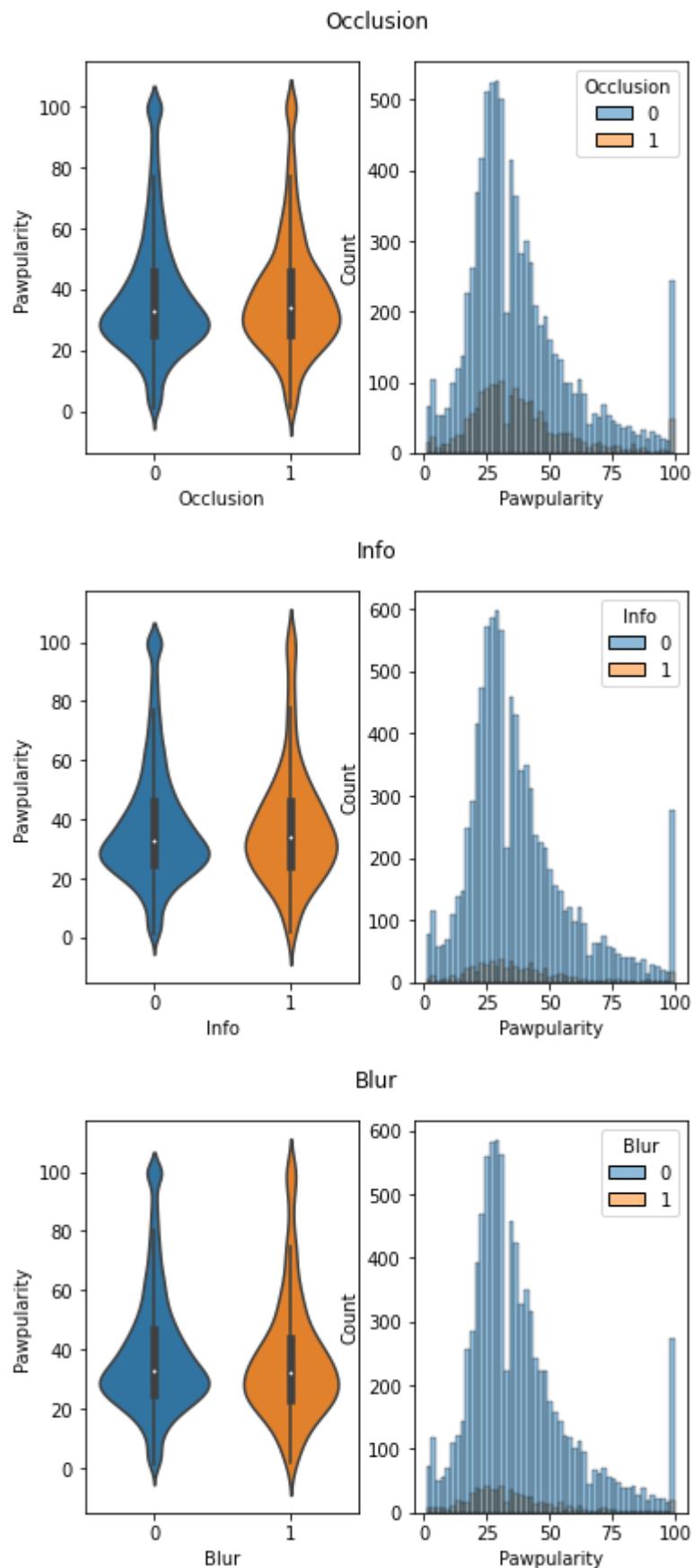
    plt.suptitle(col)
    fig.show()
```

c:\users\administrator\appdata\local\programs\python\python37\lib\site-packages\ipykernel_launcher.py:11: UserWarning: Matplotlib is currently using module://ipykernel.pylab.backend_inline, which is a non-GUI backend, so cannot show the figure.





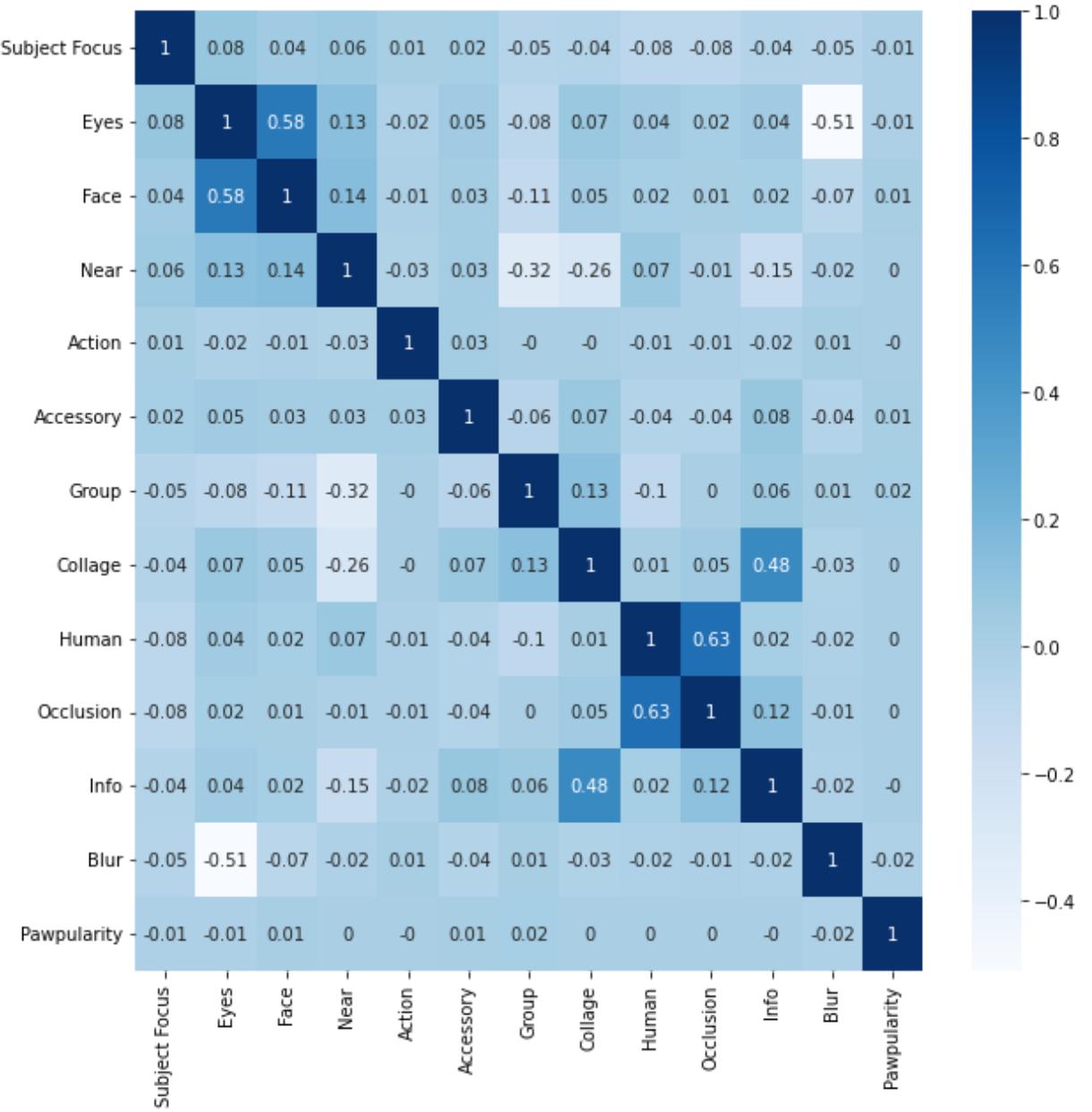




plot heatmap of data

In [8]:

```
import seaborn as sns
plt.figure(figsize=(10,10))
sns.heatmap(round(dftrainshow.corr(), 2), annot=True, cmap = 'Blues');
```



It also looks as though multicollinearity will not be a huge issue with our data

However, these characteristics were previously correlated:

Face and Eyes

Human and Occlusion

Collage and Info

Blur and Eyes

Since the Eyes-feature is highly correlated with two other features, we might consider dropping it as opposed to dropping the Face- and Blur-features.

2. Analysis of metadata

Based on the previous visualization, it is regretful that we cannot find some correlation between the metadata in each column to the Pawpularity score. However, before feeding the images into neural network, we want to make extra experiments on the metadata to see if there exist some determining factor by some traditional regression methods and feature selections, which might help the further analysis on image data.

```
In [10]: ## Extract metadata
metadata = train_df.drop(labels = 'Id', axis=1)
metadata = metadata.drop(labels = 'Pawpularity', axis=1)
metadata
```

Out[10]:

	Subject Focus	Eyes	Face	Near	Action	Accessory	Group	Collage	Human	Occlusion	Info
0	0	1	1	1	0	0	1	0	0	0	0
1	0	1	1	0	0	0	0	0	0	0	0
2	0	1	1	1	0	0	0	0	1	1	0
3	0	1	1	1	0	0	0	0	0	0	0
4	0	0	0	1	0	0	1	0	0	0	0
...
9907	0	0	0	1	0	0	0	0	0	0	0
9908	0	1	1	1	0	0	0	0	0	0	0
9909	0	1	1	1	0	0	0	0	1	1	0
9910	0	1	1	1	0	0	0	0	1	0	0
9911	0	1	1	1	0	0	0	0	0	0	0

9912 rows × 12 columns

```
In [19]: import numpy as np
X_raw = metadata.values
Y = train_df['Pawpularity'].values
print(X_raw.shape)
print(Y.shape)
```

(9912, 12)
(9912,)

```
In [21]: from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(X_raw, Y, test_size=0.2, )
print(x_train.shape)
print(x_test.shape)
print(y_train.shape)
print(y_test.shape)
```

(7929, 12)
(1983, 12)
(7929,)
(1983,)

2.1 Ordinary Least Squares

```
In [26]: ols = linear_model.LinearRegression()
ols.fit(x_train, y_train)
```

Out[26]: LinearRegression()

```
In [35]: def eval_predict(trueY, predY):
    RMSE = np.sqrt(metrics.mean_squared_error(trueY, predY))
    return RMSE
```

```

def plot_scatter(testY, predY, titlestr):
    plt.scatter(testY, predY, s=1)
    plt.plot([min(testY),max(testY)], [min(testY),max(testY)], 'r-')
    plt.xlabel('true')
    plt.ylabel('prediction')
    r1,r2 = eval_predict(testY, predY)
    plt.title(titlestr + "\n(RMSE={:.3f}; MAE={:.3f})".format(r1,r2))
    plt.grid(True)

```

```

In [34]: trainerr={}; testerr={}
trainerr['ols'] = eval_predict(y_train, ols.predict(x_train))
testerr['ols'] = eval_predict(y_test, ols.predict(x_test))
print("OLS: train RMSE =", trainerr['ols'])
print("OLS: test RMSE =", testerr['ols'])

```

```

OLS: train RMSE = 20.50251846689939
OLS: test RMSE = 20.827033388978084

```

```

In [43]: print(ols.predict(x_train))

```

```

[37.9593998 37.9593998 38.24662603 ... 39.00192313 38.02751381
 37.9593998 ]

```

2.1 Ridge Regression

```

In [44]: alphas = np.logspace(-3,3,30)
rr = linear_model.RidgeCV(alphas=alphas, cv=5)
rr.fit(x_train, y_train)
print("alpha =", rr.alpha_)
print("w =", rr.coef_)

alpha = 1000.0
w = [-0.31969042 -0.30862504  0.20862457  0.25568166 -0.07541523  0.31909741
 0.41862805 -0.08064604  0.10952803  0.0629402  -0.15547713 -0.71223707]

```

```

In [46]: trainerr['rr'] = eval_predict(y_train, rr.predict(x_train))
testerr['rr'] = eval_predict(y_test, rr.predict(x_test))
print("RR: train RMSE =", trainerr['rr'])
print("RR: test RMSE =", testerr['rr'])

```

```

RR: train RMSE = 20.514597504565458
RR: test RMSE = 20.850728451054874

```

```

In [47]: print(rr.predict(x_train))

```

```

[38.01222985 38.01222985 38.18469807 ... 38.4308579 38.07517005
 38.01222985]

```

2.3 LASSO

```

In [48]: las = linear_model.LassoCV()
las.fit(x_train, y_train)

```

```

Out[48]: LassoCV()

```

```

In [49]: trainerr['las'] = eval_predict(y_train, las.predict(x_train))
testerr['las'] = eval_predict(y_test, las.predict(x_test))
print("LAS: train RMSE =", trainerr['las'])
print("LAS: test RMSE =", testerr['las'])

```

```
LAS: train RMSE = 20.522960427499495
LAS: test RMSE = 20.860644352486133
```

```
In [50]: print(las.predict(x_train))
```

```
[38.0587716 38.0587716 38.0587716 ... 38.0587716 38.0587716 38.0587716]
```

2.4 OMP

```
In [51]: omp = linear_model.OrthogonalMatchingPursuit(n_nonzero_coefs=5)
omp.fit(x_train, y_train)
```

```
Out[51]: OrthogonalMatchingPursuit(n_nonzero_coefs=5)
```

```
In [52]: trainerr['omp'] = eval_predict(y_train, omp.predict(x_train))
testerr['omp'] = eval_predict(y_test, omp.predict(x_test))
print("OMP: train RMSE =", trainerr['omp'])
print("OMP: test RMSE =", testerr['omp'])
```

```
OMP: train RMSE = 20.506505678302045
OMP: test RMSE = 20.825854736368807
```

```
In [57]: print(omp.predict(x_train))
```

```
[37.99163503 37.99163503 37.99163503 ... 38.70310142 37.99163503
 37.99163503]
```

The result of trial on regression methods is similar to observation on the visualization of metadata---It's obvious that traditional regression methods(neither feature selected or not) could only predict a result around the mean value of Parpularity score. There is no such determining factor to predict the Parpularity score directly.

Hence we know it's impossible to ignore the image data. Since the traditional regression could be treated as a dunmmy regression, we also get another information that the baseline of RMSE is around 20-21. In the training of neural network, we will try to conquer this baseline to get a lower RMSE.

2.5 Dimensionality reduction

Another analysis of metadata is to try to perform dimensionality reduction. Here we reduce the dim of 12-column metadata to 2-dim to see if there is new discovery.

```
In [80]: #change train_y to 20-Decile
import math
def transform_y(value):
    return math.ceil(value/5)*5

tmp = y_train.tolist()
tmp2 = [transform_y(i) for i in tmp]
y_train_20 = np.array(tmp2)
```

```
In [82]: print(y_train_20)
print(y_train)
```

```
[30 15 15 ... 30 35 70]
[26 11 15 ... 29 31 68]
```

```
In [87]: pca = decomposition.PCA(n_components=2)
W = pca.fit_transform(x_train)
print(pca.components_)
```

```
[[ -1.07760504e-02 -6.70237514e-01 -3.52995962e-01 -2.22966356e-01
   4.87523821e-03 -2.32023716e-02  1.95515968e-01 -1.00309698e-02
  -3.90226199e-01 -3.68164646e-01 -2.65869880e-02  2.21195870e-01
 [-4.72628409e-02 -4.45113660e-01 -2.39153921e-01 -1.01582311e-01
  -9.76787891e-05 -5.28998141e-02  4.80980515e-02  1.39014923e-02
  5.77722191e-01  6.07315560e-01  3.47982081e-02  1.51708211e-01]]
```

```
In [ ]: plt.figure()
plt.scatter(W[:,0], W[:,1], c=y_train_20, cmap=rbow, edgecolors='k')
plt.xlabel('$w_1$'); plt.ylabel('$w_2$')
plt.title('PCA'); plt.grid(True)
```

3. CNN

```
In [ ]: # setup
%matplotlib inline
import IPython.core.display
IPython.core.display.set_matplotlib_formats("svg")
import matplotlib.pyplot as plt
import matplotlib
from numpy import *
from sklearn import *
from scipy import stats
from scipy import signal

rbow = plt.get_cmap('rainbow')
```

```
In [3]: import tensorflow as tf
from tensorflow import keras
from tensorflow.keras.models import Sequential, Model
from tensorflow.keras.layers import Dense, Activation, Conv2D, Flatten, Dropout
from tensorflow.keras import backend as K
import logging
logging.basicConfig()
import struct
print(keras.__version__, tf.__version__)
K.set_image_data_format('channels_last')
```

2.4.0 2.4.0

```
In [4]: import pandas as pd
filepath = 'petfinder-pawpularity-score/'
train_df = pd.read_csv(filepath + 'train.csv')
train_df
```

```
Out[4]:
```

	Id	Subject Focus	Eyes	Face	Near	Action	Accessory	Gift
0	0007de18844b0dbbb5e1f607da0606e0	0	1	1	1	0	0	0
1	0009c66b9439883ba2750fb825e1d7db	0	1	1	0	0	0	0
2	0013fd999caf9a3efe1352ca1b0d937e	0	1	1	1	0	0	0
3	0018df346ac9c1d8413cfcc888ca8246	0	1	1	1	0	0	0

			Id	Subject Focus	Eyes	Face	Near	Action	Accessory	Group
4	001dc955e10590d3ca4673f034feef2				0	0	0	1	0	0
...
9907	ffbfa0383c34dc513c95560d6e1fdb57				0	0	0	1	0	0
9908	ffcc8532d76436fc79e50eb2e5238e45				0	1	1	1	0	0
9909	ffdf2e8673a1da6fb80342fa3b119a20				0	1	1	1	0	0
9910	fff19e2ce11718548fa1c5d039a5192a				0	1	1	1	0	0
9911	fff8e47c766799c9e12f3cb3d66ad228				0	1	1	1	0	0

9912 rows × 14 columns

```
In [4]: test_df = pd.read_csv(filepath + 'test.csv')
test_df
```

		Id	Subject Focus	Eyes	Face	Near	Action	Accessory	Group	
0	4128bae22183829d2b5fea10effdb0c3			1	0	1	0	0	1	1
1	43a2262d7738e3d420d453815151079e			0	1	0	0	0	0	1
2	4e429cead1848a298432a0acad014c9d			0	0	0	1	0	1	1
3	80bc3ccafcc51b66303c2c263aa38486			1	0	1	0	0	0	0
4	8f49844c382931444e68dffbe20228f4			1	1	1	0	1	1	0
5	b03f7041962238a7c9d6537e22f9b017			0	0	1	1	1	1	1
6	c978013571258ed6d4637f6e8cc9d6a3			1	0	0	0	1	1	0
7	e0de453c1bffc20c22b072b34b54e50f			1	0	1	0	0	0	0

As we learned from lecture note, CNN is suitable for this task due to its good feature extracting ability.

In this part, we will try CNN on the images and calculate the Pawpularity score.

Load data

```
In [7]: %%time
import skimage.io
import skimage.transform
img_file_names = []
imgdata = []
classes = []
for i in range(len(train_df['Id'].tolist())):
    img_file_name = filepath + 'train/' + train_df['Id'].tolist()[i] + '.jpg'
    img_file_names.append(img_file_name)
    img = skimage.io.imread(img_file_name)
    img2 = skimage.transform.resize(img, (128,128), anti_aliasing=False)
    imgdata.append(img2)
    classes.append(train_df['Pawpularity'].tolist()[i])
```

Wall time: 5min 21s

```
In [9]: %%time
plt.figure(figsize=(9,5))
for i in range(8):
    plt.subplot(2,4,i+1)
    ind = i*100
    plt.imshow(imgdata[ind])
    plt.title('Pawpularity: '+str(classes[ind]))
    plt.axis('off')
plt.show()
```

Pawpularity: 63



Pawpularity: 75



Pawpularity: 26



Pawpularity: 53



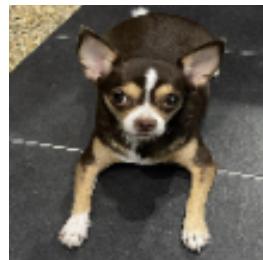
Pawpularity: 24



Pawpularity: 44



Pawpularity: 55



Pawpularity: 84



Wall time: 408 ms

```
In [8]: imgdata[0].shape
```

```
Out[8]: (128, 128, 3)
```

```
In [10]:
```

```
%%time
import numpy as np
from keras.preprocessing import image
import tensorflow.keras.applications.resnet50 as resnet
Xim_list=[]
for img in imgdata:
    xi = image.img_to_array(img)
    #xi = expand_dims(xi, axis=0)
    #xi = resnet.preprocess_input(xi)
    Xim_list.append(xi)
Xim = np.array(Xim_list)
print(Xim.shape)
```

```
(9912, 128, 128, 3)
```

Wall time: 38.5 s

```
In [12]:
```

```
from sklearn.model_selection import train_test_split
classes = np.array(classes)
x_train, x_test, y_train, y_test = train_test_split(Xim, classes, test_size=0
print(x_train.shape)
print(x_test.shape)
print(y_train.shape)
print(y_test.shape)
```

```
(6938, 128, 128, 3)
(2974, 128, 128, 3)
(6938,)
(2974,)
```

We will use 70% of data in train set as training data and another 30% to do monitoring.

3.1 CNN Model1

As we know, one advantage of Convolution kernel of CNN is to "identify" some structures on the images. And some structures on this task such as how large is the pet's face, is there some undesirable objects.....will affect the Pawpularity score, which is what we need to predict on this task.

This CNN model was learned from online resources and have a not bad performance on this task.

Because our input shape is [128,128,3], the first Convolution kernel with size [7,7] and step 2 might be able to extract some "important points" on pet images such as faces and special points of background. For the deeper layers, it is hard to deduct their effect now and we will try to do a visualization and analysis of the feature map after training this model.

```
In [13]: from tensorflow.keras.regularizers import l2
def get_model():
    tf.keras.backend.clear_session()
    np.random.seed(4487)
    tf.random.set_seed(4487)

    input_ = keras.Input(shape=(128,128,3))
    x = keras.layers.Conv2D(filters = 16, kernel_size=[7,7], strides=[2,2], padding='same', kernel_regularizer=l2(0.0005), activation = 'relu')(x)
    x = keras.layers.Conv2D(filters = 32, kernel_size=[3,3], padding='same', activation = 'relu')(x)
    x = keras.layers.BatchNormalization()(x)
    x = keras.layers.Conv2D(filters = 32, kernel_size=[3,3], strides=[2,2], padding='same', activation = 'relu', kernel_regularizer=l2(0.0005))(x)
    x = keras.layers.BatchNormalization()(x)
    x = keras.layers.Dropout(0.25)(x)
    x = keras.layers.Conv2D(filters = 64, kernel_size=[3,3], padding='same', kernel_regularizer=l2(0.0002))(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.Conv2D(filters = 64, kernel_size = (3,3), strides = (1,1), kernel_initializer='he_normal', kernel_regularizer=l2(0.0005))(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.Dropout(0.25)(x)

    x = tf.keras.layers.Conv2D(filters = 128, kernel_size = (3,3), padding='same')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.MaxPooling2D(pool_size=(2, 2))(x)
    x = tf.keras.layers.Conv2D(filters = 128, kernel_size = (3,3), padding='same')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.Dropout(0.25)(x)
    x = tf.keras.layers.Flatten()(x)
    x = tf.keras.layers.Dense(512, activation = "relu")(x)
    x = tf.keras.layers.Dropout(0.5)(x)
    output = tf.keras.layers.Dense(1)(x)
    model = tf.keras.Model(inputs = input_, outputs = output)
    return model
```

```
In [14]: model_1 = get_model()
model_1.summary()
```

```
Model: "model"
```

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 128, 128, 3)]	0
conv2d (Conv2D)	(None, 61, 61, 16)	2368
conv2d_1 (Conv2D)	(None, 61, 61, 32)	4640
batch_normalization (BatchNo	(None, 61, 61, 32)	128
conv2d_2 (Conv2D)	(None, 31, 31, 32)	9248
batch_normalization_1 (Batch	(None, 31, 31, 32)	128
dropout (Dropout)	(None, 31, 31, 32)	0
conv2d_3 (Conv2D)	(None, 31, 31, 64)	18496
batch_normalization_2 (Batch	(None, 31, 31, 64)	256
conv2d_4 (Conv2D)	(None, 16, 16, 64)	36928
batch_normalization_3 (Batch	(None, 16, 16, 64)	256
dropout_1 (Dropout)	(None, 16, 16, 64)	0
conv2d_5 (Conv2D)	(None, 16, 16, 128)	73856
batch_normalization_4 (Batch	(None, 16, 16, 128)	512
max_pooling2d (MaxPooling2D)	(None, 8, 8, 128)	0
conv2d_6 (Conv2D)	(None, 8, 8, 128)	147584
batch_normalization_5 (Batch	(None, 8, 8, 128)	512
dropout_2 (Dropout)	(None, 8, 8, 128)	0
flatten (Flatten)	(None, 8192)	0
dense (Dense)	(None, 512)	4194816
dropout_3 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 1)	513
<hr/>		
Total params: 4,490,241		
Trainable params: 4,489,345		
Non-trainable params: 896		

Based on the requirement of kaggle competition, we use rmse as loss to train the model.

```
In [15]:
```

```
model_1.compile(  
    loss = 'mse',  
    optimizer = 'Adam',  
    metrics = [tf.keras.metrics.RootMeanSquaredError(name="rmse")])
```

Here we also apply the data augmentation to add noises to data in order to prevent overfitting.

```
In [16]:
```

```
from tensorflow.keras.preprocessing.image import ImageDataGenerator  
data_augmentation = ImageDataGenerator(rotation_range=15, zoom_range=0.15, wi
```

```
height_shift_range = 0.2,  
shear_range = 0.1,  
horizontal_flip = True,  
fill_mode = "nearest")
```

After totally 60 epochs of training, we got a val_rmse around 21.0

```
In [18]:  
earlystop = keras.callbacks.EarlyStopping(monitor='val_rmse', patience=10, restore_best_weights=True)  
  
history = model_1.fit(  
    data_augmentation.flow(x_train, y_train, batch_size=1024),  
    validation_data = (x_test, y_test),  
    steps_per_epoch = len(x_train) // 1024,  
    epochs = 60, callbacks=[earlystop]  
)  
  
Epoch 1/60  
6/6 [=====] - 83s 14s/step - loss: 1090.7504 - rmse: 32.6612 - val_loss: 1964.3889 - val_rmse: 44.3187  
Epoch 2/60  
6/6 [=====] - 82s 13s/step - loss: 529.6457 - rmse: 23.0030 - val_loss: 1433.9664 - val_rmse: 37.8646  
Epoch 3/60  
6/6 [=====] - 80s 13s/step - loss: 473.4074 - rmse: 21.7504 - val_loss: 1265.0155 - val_rmse: 35.5637  
Epoch 4/60  
6/6 [=====] - 79s 13s/step - loss: 468.5222 - rmse: 21.6381 - val_loss: 1166.1726 - val_rmse: 34.1457  
Epoch 5/60  
6/6 [=====] - 83s 13s/step - loss: 458.4779 - rmse: 21.4056 - val_loss: 1055.6084 - val_rmse: 32.4864  
Epoch 6/60  
6/6 [=====] - 81s 13s/step - loss: 458.0011 - rmse: 21.3944 - val_loss: 1189.4054 - val_rmse: 34.4842  
Epoch 7/60  
6/6 [=====] - 80s 13s/step - loss: 436.3483 - rmse: 20.8821 - val_loss: 1068.4327 - val_rmse: 32.6832  
Epoch 8/60  
6/6 [=====] - 79s 13s/step - loss: 451.1087 - rmse: 21.2316 - val_loss: 945.0459 - val_rmse: 30.7377  
Epoch 9/60  
6/6 [=====] - 87s 15s/step - loss: 436.4470 - rmse: 20.8835 - val_loss: 938.3335 - val_rmse: 30.6283  
Epoch 10/60  
6/6 [=====] - 106s 17s/step - loss: 424.9126 - rmse: 20.6033 - val_loss: 813.6575 - val_rmse: 28.5204  
Epoch 11/60  
6/6 [=====] - 96s 16s/step - loss: 434.6322 - rmse: 20.8414 - val_loss: 818.4490 - val_rmse: 28.6043  
Epoch 12/60  
6/6 [=====] - 96s 15s/step - loss: 440.3706 - rmse: 20.9777 - val_loss: 826.1010 - val_rmse: 28.7377  
Epoch 13/60  
6/6 [=====] - 79s 13s/step - loss: 449.1138 - rmse: 21.1864 - val_loss: 776.4840 - val_rmse: 27.8611  
Epoch 14/60  
6/6 [=====] - 79s 13s/step - loss: 449.4859 - rmse: 21.1948 - val_loss: 693.3292 - val_rmse: 26.3265  
Epoch 15/60  
6/6 [=====] - 80s 13s/step - loss: 442.5834 - rmse: 21.0311 - val_loss: 714.2571 - val_rmse: 26.7210  
Epoch 16/60  
6/6 [=====] - 80s 13s/step - loss: 440.6349 - rmse: 20.9848 - val_loss: 645.6077 - val_rmse: 25.4040  
Epoch 17/60  
6/6 [=====] - 82s 13s/step - loss: 447.8439 - rmse: 21.1563 - val_loss: 624.7839 - val_rmse: 24.9908
```

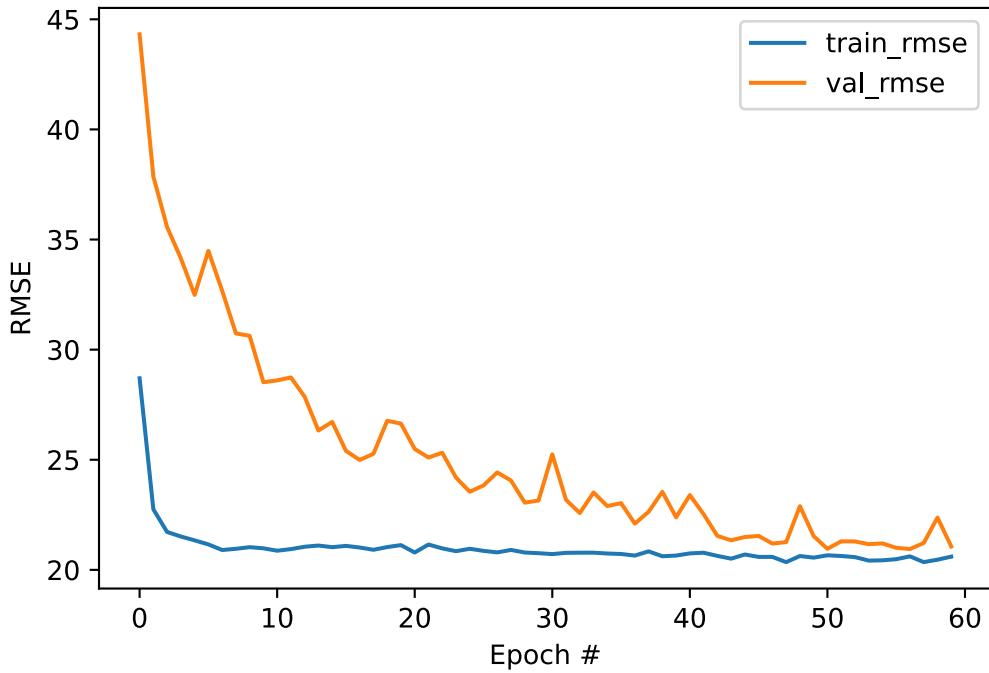
```
Epoch 18/60
6/6 [=====] - 79s 13s/step - loss: 430.2248 - rmse: 2
0.7329 - val_loss: 638.8762 - val_rmse: 25.2712
Epoch 19/60
6/6 [=====] - 79s 13s/step - loss: 440.3284 - rmse: 2
0.9777 - val_loss: 716.8995 - val_rmse: 26.7704
Epoch 20/60
6/6 [=====] - 83s 14s/step - loss: 450.6939 - rmse: 2
1.2231 - val_loss: 710.0566 - val_rmse: 26.6423
Epoch 21/60
6/6 [=====] - 81s 13s/step - loss: 423.1170 - rmse: 2
0.5624 - val_loss: 649.9573 - val_rmse: 25.4895
Epoch 22/60
6/6 [=====] - 80s 13s/step - loss: 435.0283 - rmse: 2
0.8478 - val_loss: 630.3395 - val_rmse: 25.1017
Epoch 23/60
6/6 [=====] - 78s 13s/step - loss: 440.5519 - rmse: 2
0.9834 - val_loss: 641.3080 - val_rmse: 25.3192
Epoch 24/60
6/6 [=====] - 79s 13s/step - loss: 432.6428 - rmse: 2
0.7936 - val_loss: 585.5106 - val_rmse: 24.1923
Epoch 25/60
6/6 [=====] - 85s 14s/step - loss: 450.2153 - rmse: 2
1.2106 - val_loss: 554.6966 - val_rmse: 23.5468
Epoch 26/60
6/6 [=====] - 80s 13s/step - loss: 439.9105 - rmse: 2
0.9677 - val_loss: 568.3537 - val_rmse: 23.8350
Epoch 27/60
6/6 [=====] - 79s 13s/step - loss: 431.0192 - rmse: 2
0.7545 - val_loss: 596.7753 - val_rmse: 24.4239
Epoch 28/60
6/6 [=====] - 82s 13s/step - loss: 437.6716 - rmse: 2
0.9145 - val_loss: 578.8961 - val_rmse: 24.0551
Epoch 29/60
6/6 [=====] - 79s 13s/step - loss: 432.8828 - rmse: 2
0.7994 - val_loss: 531.6843 - val_rmse: 23.0529
Epoch 30/60
6/6 [=====] - 79s 13s/step - loss: 428.0600 - rmse: 2
0.6825 - val_loss: 536.1157 - val_rmse: 23.1488
Epoch 31/60
6/6 [=====] - 79s 13s/step - loss: 424.5094 - rmse: 2
0.5973 - val_loss: 637.6473 - val_rmse: 25.2468
Epoch 32/60
6/6 [=====] - 80s 13s/step - loss: 430.6628 - rmse: 2
0.7459 - val_loss: 537.3663 - val_rmse: 23.1758
Epoch 33/60
6/6 [=====] - 80s 13s/step - loss: 422.4122 - rmse: 2
0.5452 - val_loss: 510.1204 - val_rmse: 22.5804
Epoch 34/60
6/6 [=====] - 78s 13s/step - loss: 433.1153 - rmse: 2
0.8052 - val_loss: 553.1526 - val_rmse: 23.5139
Epoch 35/60
6/6 [=====] - 79s 13s/step - loss: 443.6015 - rmse: 2
1.0541 - val_loss: 524.4705 - val_rmse: 22.8959
Epoch 36/60
6/6 [=====] - 82s 13s/step - loss: 433.9833 - rmse: 2
0.8257 - val_loss: 530.6433 - val_rmse: 23.0303
Epoch 37/60
6/6 [=====] - 79s 13s/step - loss: 423.0932 - rmse: 2
0.5629 - val_loss: 488.6591 - val_rmse: 22.1000
Epoch 38/60
6/6 [=====] - 79s 13s/step - loss: 430.4974 - rmse: 2
0.7421 - val_loss: 512.7197 - val_rmse: 22.6378
Epoch 39/60
6/6 [=====] - 79s 13s/step - loss: 431.0475 - rmse: 2
0.7551 - val_loss: 554.6572 - val_rmse: 23.5459
Epoch 40/60
6/6 [=====] - 79s 13s/step - loss: 431.3647 - rmse: 2
0.7623 - val_loss: 501.3764 - val_rmse: 22.3859
```

```
Epoch 41/60
6/6 [=====] - 82s 13s/step - loss: 439.7880 - rmse: 2
0.9611 - val_loss: 547.6313 - val_rmse: 23.3962
Epoch 42/60
6/6 [=====] - 81s 13s/step - loss: 425.8577 - rmse: 2
0.6289 - val_loss: 507.7375 - val_rmse: 22.5275
Epoch 43/60
6/6 [=====] - 80s 13s/step - loss: 415.3452 - rmse: 2
0.3729 - val_loss: 464.3029 - val_rmse: 21.5419
Epoch 44/60
6/6 [=====] - 80s 13s/step - loss: 405.6448 - rmse: 2
0.1287 - val_loss: 455.7340 - val_rmse: 21.3421
Epoch 45/60
6/6 [=====] - 82s 13s/step - loss: 421.8257 - rmse: 2
0.5309 - val_loss: 462.1944 - val_rmse: 21.4929
Epoch 46/60
6/6 [=====] - 79s 13s/step - loss: 416.1278 - rmse: 2
0.3912 - val_loss: 464.2538 - val_rmse: 21.5408
Epoch 47/60
6/6 [=====] - 79s 13s/step - loss: 417.3114 - rmse: 2
0.4208 - val_loss: 449.4642 - val_rmse: 21.1947
Epoch 48/60
6/6 [=====] - 78s 13s/step - loss: 407.1412 - rmse: 2
0.1710 - val_loss: 452.2068 - val_rmse: 21.2593
Epoch 49/60
6/6 [=====] - 80s 13s/step - loss: 418.1950 - rmse: 2
0.4425 - val_loss: 524.3304 - val_rmse: 22.8928
Epoch 50/60
6/6 [=====] - 85s 15s/step - loss: 414.6217 - rmse: 2
0.3554 - val_loss: 463.8329 - val_rmse: 21.5310
Epoch 51/60
6/6 [=====] - 91s 15s/step - loss: 419.9810 - rmse: 2
0.4863 - val_loss: 439.3809 - val_rmse: 20.9554
Epoch 52/60
6/6 [=====] - 91s 15s/step - loss: 424.2281 - rmse: 2
0.5906 - val_loss: 453.7680 - val_rmse: 21.2960
Epoch 53/60
6/6 [=====] - 92s 15s/step - loss: 425.6820 - rmse: 2
0.6257 - val_loss: 453.5671 - val_rmse: 21.2912
Epoch 54/60
6/6 [=====] - 88s 14s/step - loss: 418.9825 - rmse: 2
0.4625 - val_loss: 448.3159 - val_rmse: 21.1676
Epoch 55/60
6/6 [=====] - 92s 16s/step - loss: 420.9502 - rmse: 2
0.5103 - val_loss: 449.7114 - val_rmse: 21.2005
Epoch 56/60
6/6 [=====] - 98s 16s/step - loss: 428.1487 - rmse: 2
0.6839 - val_loss: 441.3473 - val_rmse: 21.0023
Epoch 57/60
6/6 [=====] - 80s 13s/step - loss: 416.0942 - rmse: 2
0.3914 - val_loss: 438.9724 - val_rmse: 20.9457
Epoch 58/60
6/6 [=====] - 79s 13s/step - loss: 413.3999 - rmse: 2
0.3253 - val_loss: 450.5089 - val_rmse: 21.2193
Epoch 59/60
6/6 [=====] - 79s 13s/step - loss: 429.6043 - rmse: 2
0.7178 - val_loss: 501.0121 - val_rmse: 22.3777
Epoch 60/60
6/6 [=====] - 82s 13s/step - loss: 436.7223 - rmse: 2
0.8901 - val_loss: 443.8995 - val_rmse: 21.0630
```

In [19]:

```
plt.figure()
plt.plot(history.history["rmse"], label="train_rmse")
plt.plot(history.history["val_rmse"], label="val_rmse")
plt.xlabel("Epoch #")
plt.ylabel("RMSE")
plt.legend(loc="upper right")
```

```
Out[19]: <matplotlib.legend.Legend at 0x162264bfe48>
```



```
In [27]: model_1.save('my_model1')
```

```
INFO:tensorflow:Assets written to: my_model1\assets
INFO:tensorflow:Assets written to: my_model1\assets
```

Next, we will do some analysis on the trained model to see how it predict the Pawpularity score from input image.

First let's try it on one example.

```
In [7]: reload_model_1 = tf.keras.models.load_model('my_model1')
#reload_model_1.summary()
```

```
In [49]: import skimage.io
import skimage.transform
test_img_path = filepath + 'train/' + train_df['Id'][2] + '.jpg'
test_img = skimage.io.imread(test_img_path)
test_img2 = skimage.transform.resize(test_img, (128,128), anti_aliasing=False)
tmpY = train_df['Pawpularity'][2]
plt.imshow(test_img2)
print(tmpY)
```



```
In [50]:  
import numpy as np  
from keras.preprocessing import image  
xi = image.img_to_array(test_img2)  
xi = np.expand_dims(xi, axis=0)  
print(xi.shape)  
  
(1, 128, 128, 3)
```

Here we got the predicted score is 32, which is close to the ground truth of 28.

```
In [51]:  
preds = reload_model_1.predict(xi)  
print('predicted', preds)  
  
predicted [[31.928402]]
```

Then in order to observe how the CNN model transforms input data into a prediction result, we list out all feature maps of totally 6 convolution layers.

```
In [52]:  
for layer in reload_model_1.layers:  
    print(layer.name)  
  
input_1  
conv2d  
conv2d_1  
batch_normalization  
conv2d_2  
batch_normalization_1  
dropout  
conv2d_3  
batch_normalization_2  
conv2d_4  
batch_normalization_3  
dropout_1  
conv2d_5  
batch_normalization_4  
max_pooling2d  
conv2d_6  
batch_normalization_5  
dropout_2  
flatten  
dense
```

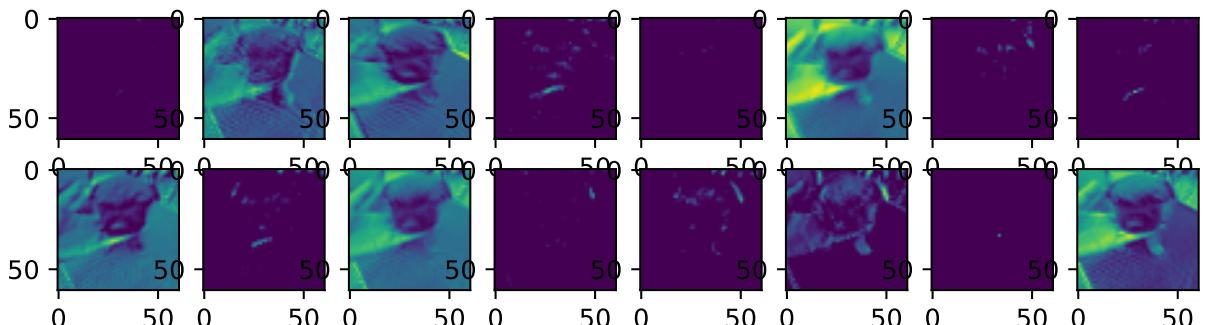
```
dropout_3  
dense_1
```

In [83]:

```
#feature maps of first convolutional layer  
layer_names = ['conv2d']  
layer_outputs = [reload_model_1.get_layer(layer_name).output for layer_name in layer_names]  
activation_model = keras.models.Model(inputs=reload_model_1.input,outputs=layer_outputs)  
activations = activation_model.predict(xi)  
plt.figure(figsize=(8,2))  
for i in range(16):  
    plt.subplot(2,8,i+1)  
    plt.imshow(activations[0,:,:,:,i],cmap='viridis')  
plt.show()
```

WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E756361F8> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your @tf.function outside of the loop. For (2), @tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling_retracing and https://www.tensorflow.org/api_docs/python/tf/function for more details.

WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E756361F8> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your @tf.function outside of the loop. For (2), @tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling_retracing and https://www.tensorflow.org/api_docs/python/tf/function for more details.



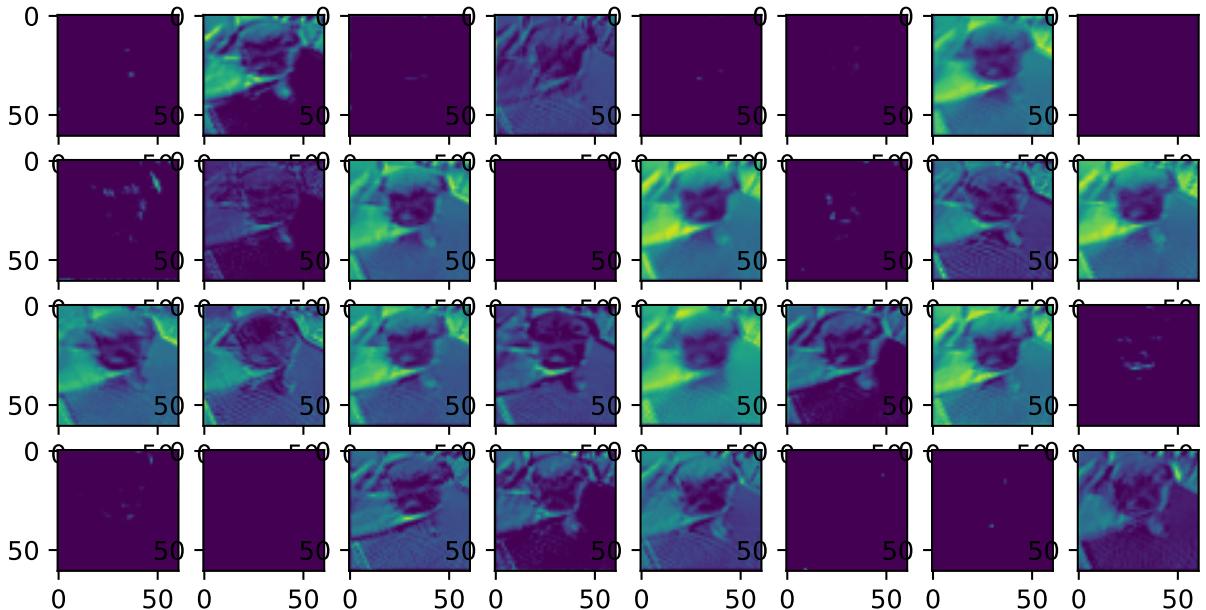
- From the feature maps of first convolution layer, we can find that some kernals (like the 6th of first row) focus on the existence of human (one guess is the kernal is sensitive to color of skin), some kernals focus on the outline of dog's face, eyes and mouth (like the 2nd of second row).

In [82]:

```
#feature maps of second convolutional layer  
layer_names = ['conv2d_1']  
layer_outputs = [reload_model_1.get_layer(layer_name).output for layer_name in layer_names]  
activation_model = keras.models.Model(inputs=reload_model_1.input,outputs=layer_outputs)  
activations = activation_model.predict(xi)  
plt.figure(figsize=(8,4))  
for i in range(32):  
    plt.subplot(4,8,i+1)  
    plt.imshow(activations[0,:,:,:,i],cmap='viridis')  
plt.show()
```

```
WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E756E7B88> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your @tf.function outside of the loop. For (2), @tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling\_retracing and https://www.tensorflow.org/api\_docs/python/tf/function for more details.
```

```
WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E756E7B88> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your @tf.function outside of the loop. For (2), @tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling\_retracing and https://www.tensorflow.org/api\_docs/python/tf/function for more details.
```



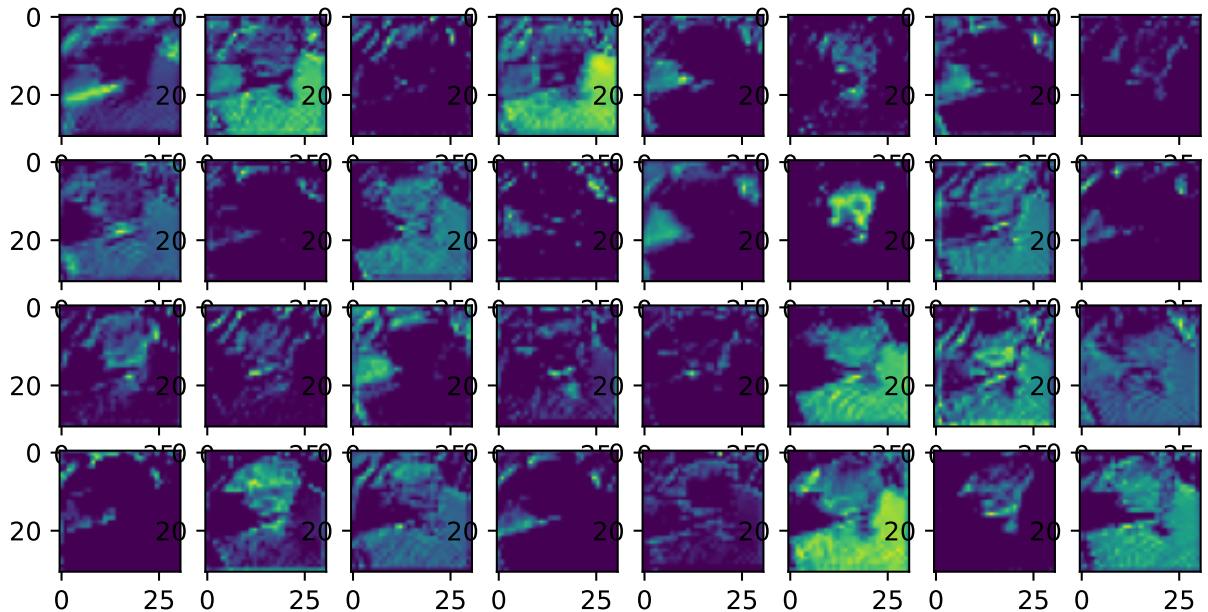
In [84]:

```
#feature maps of third convolutional layer
layer_names = ['conv2d_3']
layer_outputs = [reload_model_1.get_layer(layer_name).output for layer_name in layer_names]
activation_model = keras.models.Model(inputs=reload_model_1.input,outputs=layer_outputs)
activations = activation_model.predict(xi)
plt.figure(figsize=(8,4))
for i in range(32):
    plt.subplot(4,8,i+1)
    plt.imshow(activations[0,:,:,:,i],cmap='viridis')
plt.show()
```

```
WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E74B80708> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your @tf.function outside of the loop. For (2), @tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling\_retracing and https://www.tensorflow.org/api\_docs/python/tf/function for more details.
```

```
WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E74B80708> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could
```

be due to (1) creating `@tf.function` repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your `@tf.function` outside of the loop. For (2), `@tf.function` has `experimental_relax_shapes=True` option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling_retracing and https://www.tensorflow.org/api_docs/python/tf/function for more details.



And from the feature map after 3 convolutional layers, we can find that there are different emphasis of each kernel.

- The first feature map of first row highlights the nearest part of the human's hand, which might have correlation to the value of metadata 'human' and 'Occlusion'!
- The second feature map of last row highlights the whole area of dog's face, which will relate to the 'near' and 'face' from metadata.
- The 4th feature map of first row highlights all the background area except pet and human(or near objects?), which might relate to 'near' as well.
- ...

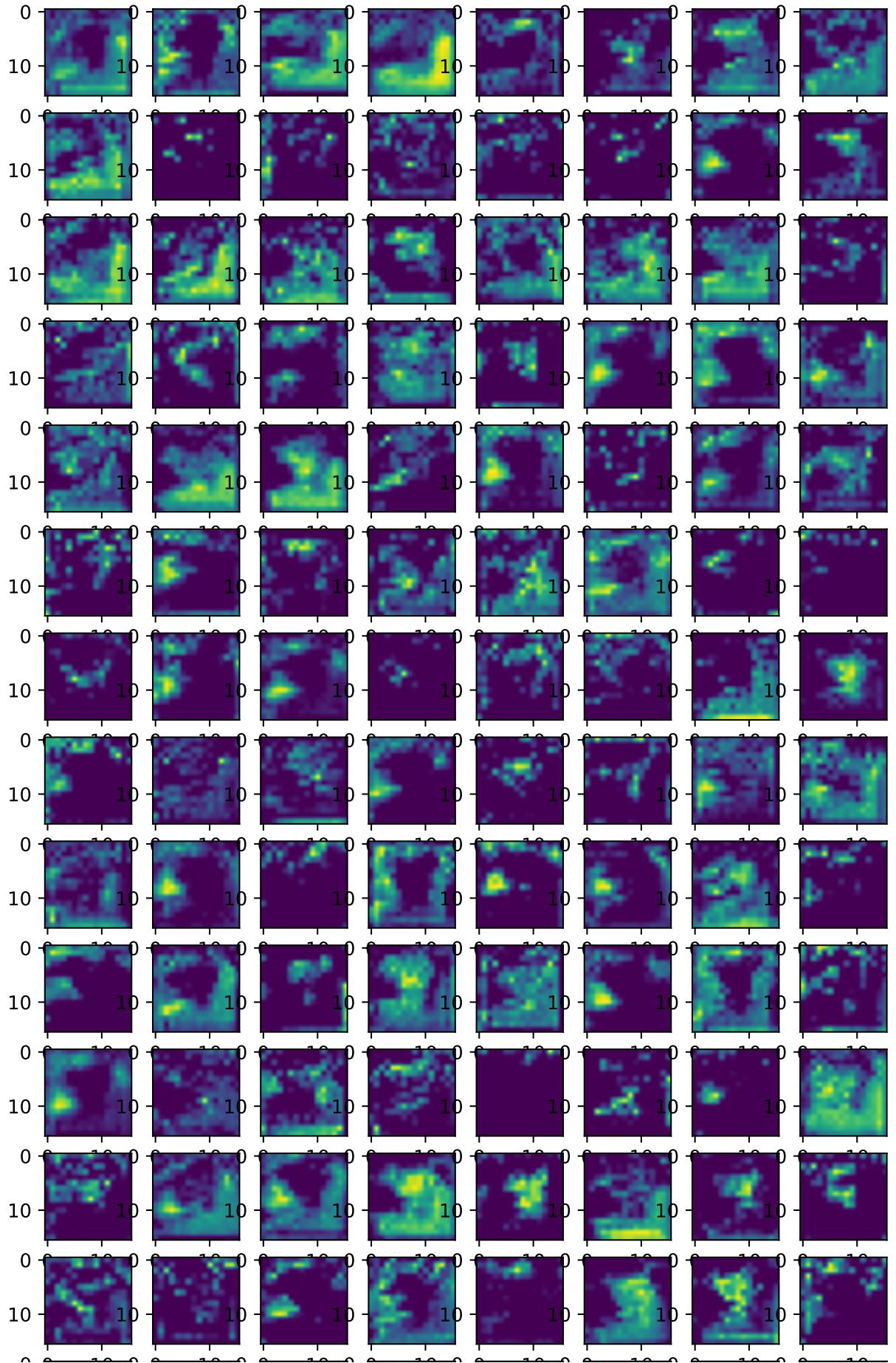
In [86]:

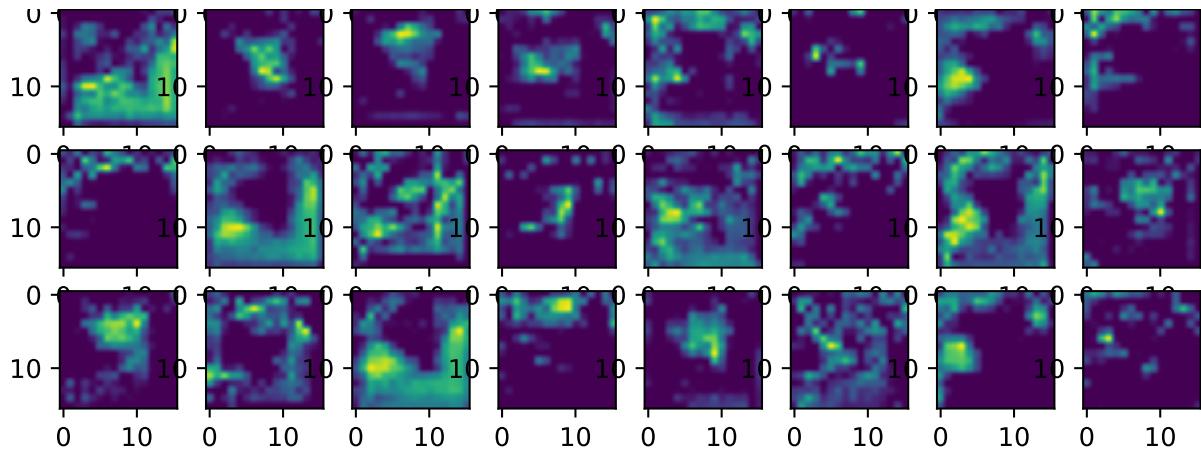
```
#feature maps of fifth convolutional layer
layer_names = ['conv2d_5']
layer_outputs = [reload_model_1.get_layer(layer_name).output for layer_name in layer_names]
activation_model = keras.models.Model(inputs=reload_model_1.input,outputs=layer_outputs)
activations = activation_model.predict(xi)
plt.figure(figsize=(8,16))
for i in range(128):
    plt.subplot(16,8,i+1)
    plt.imshow(activations[0,:,:,:,i],cmap='viridis')
plt.show()
```

WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E7FB92828> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating `@tf.function` repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your `@tf.function` outside of the loop. For (2), `@tf.function` has `experimental_relax_shapes=True` option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling_retracing and https://www.tensorflow.org/api_docs/python/tf/function for more details.

WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E7FB92828> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating `@tf.function` repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your `@tf.function` outside of the loop. For (2), `@tf.function` has `experimental_relax_shapes=True` option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling_retracing and https://www.tensorflow.org/api_docs/python/tf/function for more details.

on retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating `@tf.function` repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your `@tf.function` outside of the loop. For (2), `@tf.function` has `experimental_relax_shapes=True` option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling_retracing and https://www.tensorflow.org/api_docs/python/tf/function for more details.



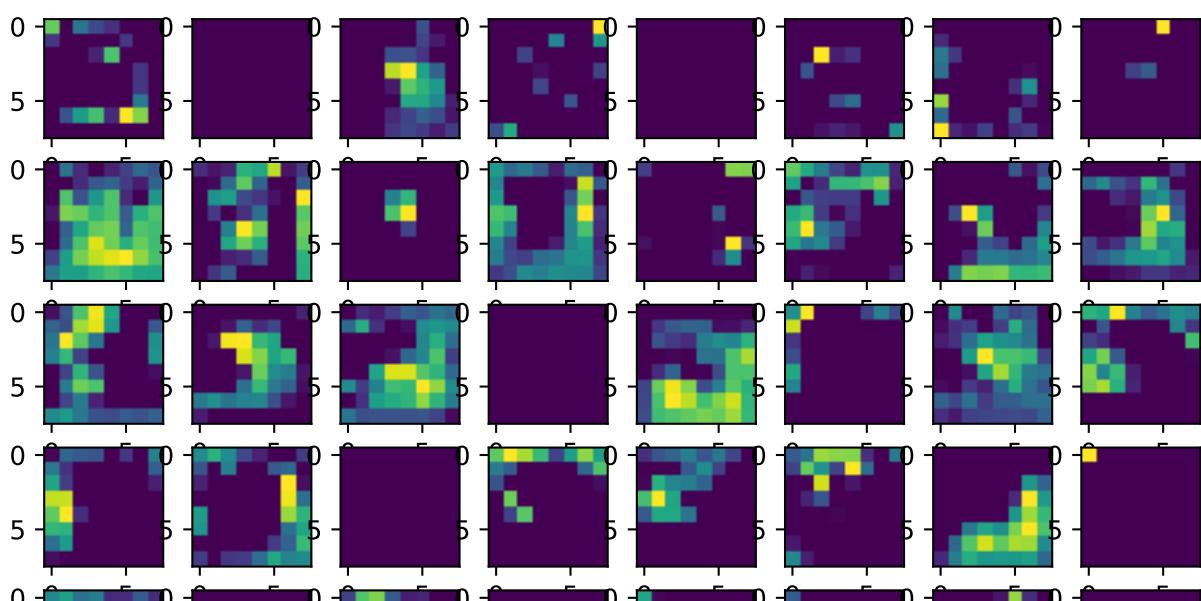


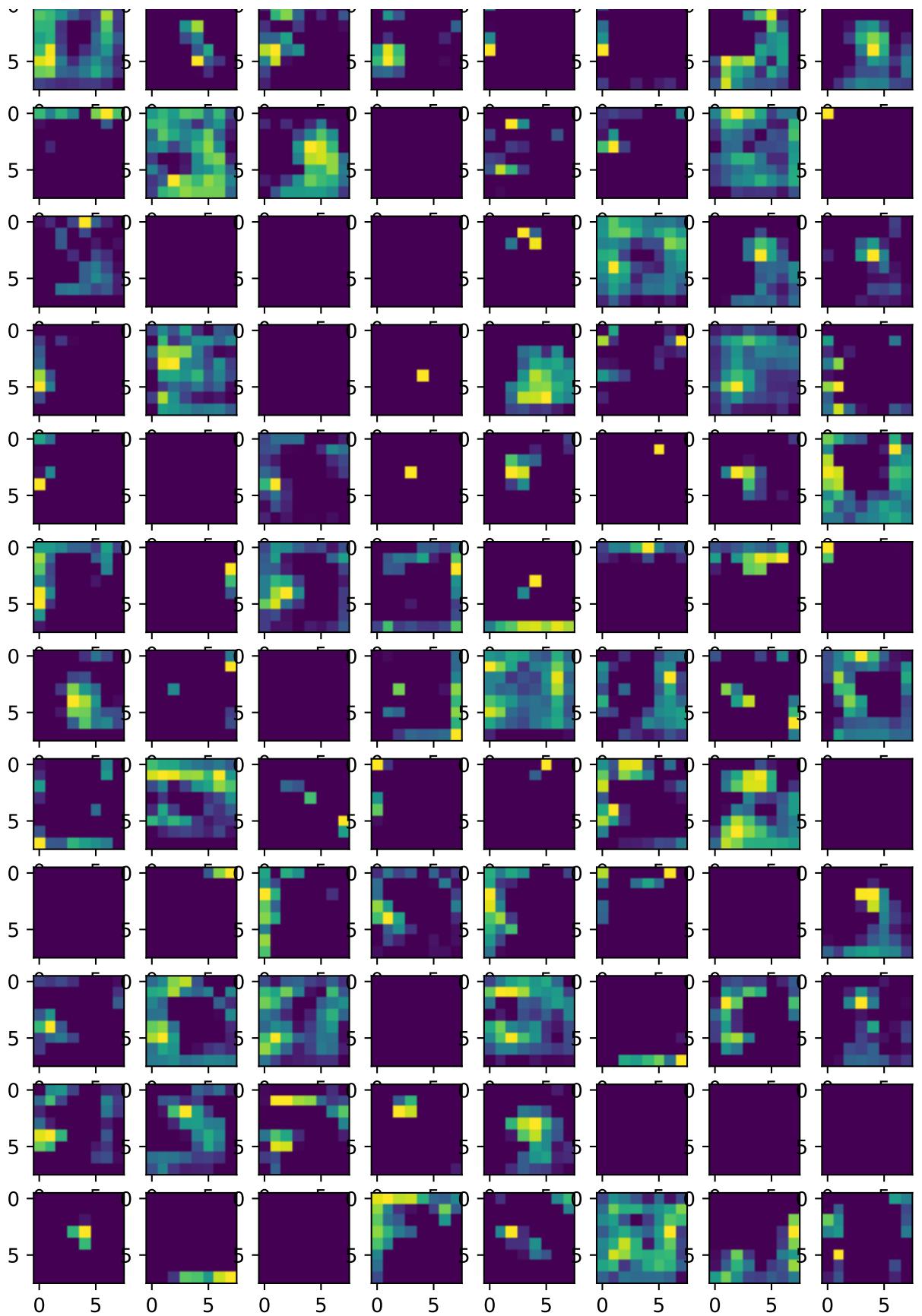
In [85]:

```
#feature maps of last convolutional layer
layer_names = ['conv2d_6']
layer_outputs = [reload_model_1.get_layer(layer_name).output for layer_name in layer_names]
activation_model = keras.models.Model(inputs=reload_model_1.input,outputs=layer_outputs)
activations = activation_model.predict(xi)
plt.figure(figsize=(8,16))
for i in range(128):
    plt.subplot(16,8,i+1)
    plt.imshow(activations[0,:,:,:,i],cmap='viridis')
plt.show()
```

WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E7A2951F8> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your @tf.function outside of the loop. For (2), @tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling_retracing and https://www.tensorflow.org/api_docs/python/tf/function for more details.

WARNING:tensorflow:11 out of the last 11 calls to <function Model.make_predict_function.<locals>.predict_function at 0x0000016E7A2951F8> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings could be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors with different shapes, (3) passing Python objects instead of tensors. For (1), please define your @tf.function outside of the loop. For (2), @tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. For (3), please refer to https://www.tensorflow.org/guide/function#controlling_retracing and https://www.tensorflow.org/api_docs/python/tf/function for more details.





The feature maps of last convolutional layer is hard to identify by human so we choose to analyze the 5th ones.

- It's obvious that we can observe some feature maps "draw a circle around the middle part". And it is reasonable to predict they are the kernels to identify the possible background structure(i.e. most pet pictures with good focus/near will put the pet's position in the middle).

- Another interesting feature map is like the 3rd of fifth row. By comparing it with the original image, we can find the extremely highlight yellow point is corresponding to the dog's nose, which means it has the ability to extract face features of pets.
- ...

3.2 CNN Model 2

EfficientNet is a Convolutional Neural Network presented by Google. It is an improvement version of ResNet. And here we will use it as the backbone network to extract feature maps from input images.

```
In [20]: from tensorflow.keras.applications import EfficientNetB5
input_picture = Input(shape=(128,128,3))
model_2 = EfficientNetB5(
    weights='pretrain_paras/efficientnetb5_notop.h5',
    include_top=False,
    input_tensor=input_picture)
model_2.trainable = False

x = GlobalAveragePooling2D()(model_2.output)
x = BatchNormalization()(x)
x = Dropout(0.1)(x)
output = Dense(32)(x)
output = Dense(1)(output)
model_2 = Model(inputs=input_picture, outputs=output)
```

```
In [21]: model_2.summary()
```

Model: "model_1"

Layer (type)	Output Shape	Param #	Connected to
input_2 (InputLayer)	[(None, 128, 128, 3) 0		
rescaling (Rescaling)	(None, 128, 128, 3) 0		input_2[0][0]
normalization (Normalization) [0]	(None, 128, 128, 3) 7		rescaling[0]
stem_conv_pad (ZeroPadding2D) [0][0]	(None, 129, 129, 3) 0		normalization[0][0]
stem_conv (Conv2D) [0][0]	(None, 64, 64, 48) 1296		stem_conv_pad[0][0]
stem_bn (BatchNormalization) [0]	(None, 64, 64, 48) 192		stem_conv[0]
stem_activation (Activation) [0][0]	(None, 64, 64, 48) 0		stem_bn[0][0]
block1a_dwconv (DepthwiseConv2D) [0][0]	(None, 64, 64, 48) 432		stem_activation[0][0]

block1a_bn (BatchNormalization) (None, 64, 64, 48)	192	block1a_dwcon v[0][0]
block1a_activation (Activation) (None, 64, 64, 48)	0	block1a_bn[0] [0]
block1a_se_squeeze (GlobalAvera (None, 48)	0	block1a_activ ation[0][0]
block1a_se_reshape (Reshape) (None, 1, 1, 48)	0	block1a_se_sq ueeze[0][0]
block1a_se_reduce (Conv2D) (None, 1, 1, 12)	588	block1a_se_re shape[0][0]
block1a_se_expand (Conv2D) (None, 1, 1, 48)	624	block1a_se_re duce[0][0]
block1a_se_excite (Multiply) (None, 64, 64, 48)	0	block1a_activ ation[0][0]
block1a_se_ex pand[0][0]		
block1a_project_conv (Conv2D) (None, 64, 64, 24)	1152	block1a_se_ex cite[0][0]
block1a_project_bn (BatchNormal (None, 64, 64, 24)	96	block1a_proje ct_conv[0][0]
block1b_dwconv (DepthwiseConv2D (None, 64, 64, 24)	216	block1a_proje ct_bn[0][0]
block1b_bn (BatchNormalization) (None, 64, 64, 24)	96	block1b_dwcon v[0][0]
block1b_activation (Activation) (None, 64, 64, 24)	0	block1b_bn[0] [0]
block1b_se_squeeze (GlobalAvera (None, 24)	0	block1b_activ ation[0][0]
block1b_se_reshape (Reshape) (None, 1, 1, 24)	0	block1b_se_sq ueeze[0][0]
block1b_se_reduce (Conv2D) (None, 1, 1, 6)	150	block1b_se_re shape[0][0]
block1b_se_expand (Conv2D) (None, 1, 1, 24)	168	block1b_se_re duce[0][0]
block1b_se_excite (Multiply) (None, 64, 64, 24)	0	block1b_activ ation[0][0]

pand[0][0]			block1b_se_ex
block1b_project_conv (Conv2D) (None, 64, 64, 24) 576	cite[0][0]		block1b_se_ex
block1b_project_bn (BatchNormal (None, 64, 64, 24) 96	cet_conv[0][0]		block1b_proje
block1b_drop (Dropout) (None, 64, 64, 24) 0	ct_bn[0][0]		block1b_proje
block1b_add (Add) (None, 64, 64, 24) 0	[0][0]		block1b_drop
block1a_project_bn[0][0]			block1a_proje
block1c_dwconv (DepthwiseConv2D (None, 64, 64, 24) 216	[0][0]		block1b_add
block1c_bn (BatchNormalization) (None, 64, 64, 24) 96	v[0][0]		block1c_dwcon
block1c_activation (Activation) (None, 64, 64, 24) 0	[0]		block1c_bn[0]
block1c_se_squeeze (GlobalAvera (None, 24) 0	ation[0][0]		block1c_activ
block1c_se_reshape (Reshape) (None, 1, 1, 24) 0	ueeze[0][0]		block1c_se_sq
block1c_se_reduce (Conv2D) (None, 1, 1, 6) 150	shape[0][0]		block1c_se_re
block1c_se_expand (Conv2D) (None, 1, 1, 24) 168	duce[0][0]		block1c_se_re
block1c_se_excite (Multiply) (None, 64, 64, 24) 0	activation[0][0]		block1c_activ
block1c_se_ex pand[0][0]			block1c_se_ex
block1c_project_conv (Conv2D) (None, 64, 64, 24) 576	cite[0][0]		block1c_se_ex
block1c_project_bn (BatchNormal (None, 64, 64, 24) 96	cet_conv[0][0]		block1c_proje
block1c_drop (Dropout) (None, 64, 64, 24) 0	ct_bn[0][0]		block1c_proje
block1c_add (Add) (None, 64, 64, 24) 0			block1c_drop

[0][0]			block1b_add
[0][0]			
block2a_expand_conv (Conv2D)	(None, 64, 64, 144)	3456	block1c_add
[0][0]			
block2a_expand_bn (BatchNormali	(None, 64, 64, 144)	576	block2a_expan
d_conv[0][0]			
block2a_expand_activation (Acti	(None, 64, 64, 144)	0	block2a_expan
d_bn[0][0]			
block2a_dwconv_pad (ZeroPadding	(None, 65, 65, 144)	0	block2a_expan
dwconv[0][0]			
block2a_dwconv (DepthwiseConv2D	(None, 32, 32, 144)	1296	block2a_dwcon
vn[0][0]			
block2a_bn (BatchNormalizatio	(None, 32, 32, 144)	576	block2a_dwcon
n[0][0]			
block2a_activation (Activation)	(None, 32, 32, 144)	0	block2a_bn[0]
[0]			
block2a_se_squeeze (GlobalAvera	(None, 144)	0	block2a_activ
ge[0][0]			
block2a_se_reshape (Reshape)	(None, 1, 1, 144)	0	block2a_se_sq
ueeze[0][0]			
block2a_se_reduce (Conv2D)	(None, 1, 1, 6)	870	block2a_se_re
shape[0][0]			
block2a_se_expand (Conv2D)	(None, 1, 1, 144)	1008	block2a_se_re
duce[0][0]			
block2a_se_excite (Multiply)	(None, 32, 32, 144)	0	block2a_activ
ation[0][0]			
block2a_se_ex pand[0][0]			block2a_se_ex
block2a_project_conv (Conv2D)	(None, 32, 32, 40)	5760	block2a_se_ex
cite[0][0]			
block2a_project_bn (BatchNormal	(None, 32, 32, 40)	160	block2a_proje
ct_conv[0][0]			
block2b_expand_conv (Conv2D)	(None, 32, 32, 240)	9600	block2a_proje
ct_bn[0][0]			
block2b_expand_bn (BatchNormali	(None, 32, 32, 240)	960	block2b_expan
d_conv[0][0]			

block2b_expand_activation (Acti (None, 32, 32, 240) 0	block2b_expand
block2b_dwconv (DepthwiseConv2D (None, 32, 32, 240) 2160	block2b_expand
block2b_bn (BatchNormalization) (None, 32, 32, 240) 960	block2b_dwcon
block2b_activation (Activation) (None, 32, 32, 240) 0	block2b_bn[0]
block2b_se_squeeze (GlobalAvera (None, 240) 0	block2b_activ
block2b_se_reshape (Reshape) (None, 1, 1, 240) 0	block2b_se_sq
block2b_se_reduce (Conv2D) (None, 1, 1, 10) 2410	block2b_se_re
block2b_se_expand (Conv2D) (None, 1, 1, 240) 2640	block2b_se_re
block2b_se_excite (Multiply) (None, 32, 32, 240) 0	block2b_activ
block2b_se_ex pand[0][0]	block2b_se_ex
block2b_project_conv (Conv2D) (None, 32, 32, 40) 9600	block2b_se_ex
block2b_project_bn (BatchNormal (None, 32, 32, 40) 160	block2b_proje
ct_conv[0][0]	ct
block2b_drop (Dropout) (None, 32, 32, 40) 0	block2b_proje
ct_bn[0][0]	ct
block2b_add (Add) (None, 32, 32, 40) 0	block2b_drop
block2a_project_bn[0][0]	block2a_proje
block2c_expand_conv (Conv2D) (None, 32, 32, 240) 9600	block2b_add
block2c_expand_bn (BatchNormali (None, 32, 32, 240) 960	block2c_expan
block2c_expand_conv[0][0]	block2c_expan
block2c_expand_activation (Acti (None, 32, 32, 240) 0	block2c_expan
block2c_bn[0][0]	block2c_expan

block2c_dwconv (DepthwiseConv2D (None, 32, 32, 240) 2160	block2c_expansion[0][0]
block2c_bn (BatchNormalization) (None, 32, 32, 240) 960	block2c_dwconv[0][0]
block2c_activation (Activation) (None, 32, 32, 240) 0	block2c_bn[0]
block2c_se_squeeze (GlobalAverage (None, 240) 0	block2c_activation[0][0]
block2c_se_reshape (Reshape) (None, 1, 1, 240) 0	block2c_se_squeeze[0][0]
block2c_se_reduce (Conv2D) (None, 1, 1, 10) 2410	block2c_se_reshape[0][0]
block2c_se_expand (Conv2D) (None, 1, 1, 240) 2640	block2c_se_reduce[0][0]
block2c_se_excite (Multiply) (None, 32, 32, 240) 0	block2c_activation[0][0]
block2c_se_excite (Multiply) (None, 32, 32, 240) 0	block2c_se_expand[0][0]
block2c_project_conv (Conv2D) (None, 32, 32, 40) 9600	block2c_se_excite[0][0]
block2c_project_bn (BatchNormal (None, 32, 32, 40) 160	block2c_project_conv[0][0]
block2c_drop (Dropout) (None, 32, 32, 40) 0	block2c_project_bn[0][0]
block2c_add (Add) (None, 32, 32, 40) 0	block2c_drop[0][0]
block2c_add (Add) (None, 32, 32, 40) 0	block2b_add[0][0]
block2d_expand_conv (Conv2D) (None, 32, 32, 240) 9600	block2c_add[0][0]
block2d_expand_bn (BatchNormali (None, 32, 32, 240) 960	block2d_expand_conv[0][0]
block2d_expand_activation (Acti (None, 32, 32, 240) 0	block2d_expand_bn[0][0]
block2d_dwconv (DepthwiseConv2D (None, 32, 32, 240) 2160	block2d_expand_activation[0][0]

block2d_bn (BatchNormalization) (None, 32, 32, 240) 960	v[0][0]	block2d_dwcon
block2d_activation (Activation) (None, 32, 32, 240) 0	[0]	block2d_bn[0]
block2d_se_squeeze (GlobalAvera (None, 240) 0	activation[0][0]	block2d_activ
block2d_se_reshape (Reshape) (None, 1, 1, 240) 0	ueeze[0][0]	block2d_se_sq
block2d_se_reduce (Conv2D) (None, 1, 1, 10) 2410	shape[0][0]	block2d_se_re
block2d_se_expand (Conv2D) (None, 1, 1, 240) 2640	duce[0][0]	block2d_se_re
block2d_se_excite (Multiply) (None, 32, 32, 240) 0	activation[0][0]	block2d_activ
block2d_se_excite (Multiply) (None, 32, 32, 240) 0	block2d_se_ex	pand[0][0]
block2d_project_conv (Conv2D) (None, 32, 32, 40) 9600	cite[0][0]	block2d_se_ex
block2d_project_bn (BatchNormal (None, 32, 32, 40) 160	conv[0][0]	block2d_proje
block2d_drop (Dropout) (None, 32, 32, 40) 0	ct_bn[0][0]	block2d_proje
block2d_add (Add) (None, 32, 32, 40) 0	[0][0]	block2d_drop
block2d_add (Add) (None, 32, 32, 40) 0	[0][0]	block2c_add
block2e_expand_conv (Conv2D) (None, 32, 32, 240) 9600	[0][0]	block2d_add
block2e_expand_bn (BatchNormali (None, 32, 32, 240) 960	conv[0][0]	block2e_expan
block2e_expand_activation (Acti (None, 32, 32, 240) 0	d_bn[0][0]	block2e_expan
block2e_dwconv (DepthwiseConv2D (None, 32, 32, 240) 2160	activation[0][0]	block2e_expan
block2e_bn (BatchNormalizat (None, 32, 32, 240) 960	v[0][0]	block2e_dwcon
block2e_activation (Activation) (None, 32, 32, 240) 0		block2e_bn[0]

[0]

block2e_se_squeeze (GlobalAvera (None, 240)	0	block2e_activation[0][0]
block2e_se_reshape (Reshape) (None, 1, 1, 240)	0	block2e_se_sq ueeze[0][0]
block2e_se_reduce (Conv2D) (None, 1, 1, 10)	2410	block2e_se_re duce[0][0]
block2e_se_expand (Conv2D) (None, 1, 1, 240)	2640	block2e_se_re duce[0][0]
block2e_se_excite (Multiply) (None, 32, 32, 240)	0	block2e_activ ation[0][0]
block2e_se_ex pand[0][0]		block2e_se_ex pand[0][0]
block2e_project_conv (Conv2D) (None, 32, 32, 40)	9600	block2e_se_ex cite[0][0]
block2e_project_bn (BatchNormal (None, 32, 32, 40)	160	block2e_proje ct_conv[0][0]
block2e_drop (Dropout) (None, 32, 32, 40)	0	block2e_proje ct_bn[0][0]
block2e_add (Add) (None, 32, 32, 40)	0	block2e_drop [0][0]
block2d_add [0][0]		block2e_add [0][0]
block3a_expand_conv (Conv2D) (None, 32, 32, 240)	9600	block2e_add [0][0]
block3a_expand_bn (BatchNormali (None, 32, 32, 240)	960	block3a_expan d_conv[0][0]
block3a_expand_activation (Acti (None, 32, 32, 240)	0	block3a_expan d_bn[0][0]
block3a_dwconv_pad (ZeroPadding (None, 35, 35, 240)	0	block3a_expan d_activation[0][0]
block3a_dwconv (DepthwiseConv2D (None, 16, 16, 240)	6000	block3a_dwcon v_pad[0][0]
block3a_bn (BatchNormalization) (None, 16, 16, 240)	960	block3a_dwcon v[0][0]
block3a_activation (Activation) (None, 16, 16, 240)	0	block3a_bn[0] [0]

block3a_se_squeeze (GlobalAvera (None, 240)	0	block3a_activation[0][0]
block3a_se_reshape (Reshape) (None, 1, 1, 240)	0	block3a_se_sq ueeze[0][0]
block3a_se_reduce (Conv2D) (None, 1, 1, 10)	2410	block3a_se_re shape[0][0]
block3a_se_expand (Conv2D) (None, 1, 1, 240)	2640	block3a_se_re duce[0][0]
block3a_se_excite (Multiply) (None, 16, 16, 240)	0	block3a_activation[0][0]
block3a_se_excite[0][0]		block3a_se_ex pand[0][0]
block3a_project_conv (Conv2D) (None, 16, 16, 64)	15360	block3a_se_ex cite[0][0]
block3a_project_bn (BatchNormal (None, 16, 16, 64)	256	block3a_proje ct_conv[0][0]
block3b_expand_conv (Conv2D) (None, 16, 16, 384)	24576	block3a_proje ct_bn[0][0]
block3b_expand_bn (BatchNormali (None, 16, 16, 384)	1536	block3b_expan d_conv[0][0]
block3b_expand_activation (Acti (None, 16, 16, 384)	0	block3b_expan d_bn[0][0]
block3b_dwconv (DepthwiseConv2D (None, 16, 16, 384)	9600	block3b_expan d_activation[0][0]
block3b_bn (BatchNormalization) (None, 16, 16, 384)	1536	block3b_dwcon v[0][0]
block3b_activation (Activation) (None, 16, 16, 384)	0	block3b_bn[0] [0]
block3b_se_squeeze (GlobalAvera (None, 384)	0	block3b_activ ation[0][0]
block3b_se_reshape (Reshape) (None, 1, 1, 384)	0	block3b_se_sq ueeze[0][0]
block3b_se_reduce (Conv2D) (None, 1, 1, 16)	6160	block3b_se_re shape[0][0]
block3b_se_expand (Conv2D) (None, 1, 1, 384)	6528	block3b_se_re

```
duce[0][0]
```

block3b_se_excite (Multiply)	(None, 16, 16, 384)	0	block3b_activation[0][0]
block3b_se_ex pand[0][0]			block3b_se_ex
block3b_project_conv (Conv2D)	(None, 16, 16, 64)	24576	block3b_se_ex cite[0][0]
block3b_project_bn (BatchNormal	(None, 16, 16, 64)	256	block3b_proje ct_conv[0][0]
ct_bn[0][0]			block3b_proje
ct_drop (Dropout)	(None, 16, 16, 64)	0	block3b_proje ct_bn[0][0]
ct_add (Add)	(None, 16, 16, 64)	0	block3b_drop [0][0]
ct_bn[0][0]			block3a_proje
ct_bn[0][0]			ct_bn[0][0]
block3c_expand_conv (Conv2D)	(None, 16, 16, 384)	24576	block3b_add [0][0]
block3c_expand_bn (BatchNormali	(None, 16, 16, 384)	1536	block3c_expan d_conv[0][0]
block3c_expand_activation (Acti	(None, 16, 16, 384)	0	block3c_expan d_bn[0][0]
block3c_dwconv (DepthwiseConv2D)	(None, 16, 16, 384)	9600	block3c_expan d_activation[0][0]
block3c_bn (BatchNormalizatio	(None, 16, 16, 384)	1536	block3c_dwcon v[0][0]
block3c_activation (Activation)	(None, 16, 16, 384)	0	block3c_bn[0][0]
block3c_se_squeeze (GlobalAvera	(None, 384)	0	block3c_activ ation[0][0]
block3c_se_reshape (Reshape)	(None, 1, 1, 384)	0	block3c_se_sq ueeze[0][0]
block3c_se_reduce (Conv2D)	(None, 1, 1, 16)	6160	block3c_se_re shape[0][0]
block3c_se_expand (Conv2D)	(None, 1, 1, 384)	6528	block3c_se_re duce[0][0]
block3c_se_excite (Multiply)	(None, 16, 16, 384)	0	block3c_activ ation[0][0]

block3c_se_ex				
pand[0][0]				
block3c_project_conv (Conv2D) (None, 16, 16, 64) 24576	block3c_se_ex			
cite[0][0]				
block3c_project_bn (BatchNormal (None, 16, 16, 64) 256	block3c_proje			
ct_conv[0][0]				
block3c_drop (Dropout) (None, 16, 16, 64) 0	block3c_proje			
ct_bn[0][0]				
block3c_add (Add) (None, 16, 16, 64) 0	block3c_drop			
[0][0]				
block3b_add				
[0][0]				
block3d_expand_conv (Conv2D) (None, 16, 16, 384) 24576	block3c_add			
[0][0]				
block3d_expand_bn (BatchNormali (None, 16, 16, 384) 1536	block3d_expan			
d_conv[0][0]				
block3d_expand_activation (Acti (None, 16, 16, 384) 0	block3d_expan			
d_bn[0][0]				
block3d_dwconv (DepthwiseConv2D (None, 16, 16, 384) 9600	block3d_expan			
d_activation[0][0]				
block3d_bn (BatchNormalizati (None, 16, 16, 384) 1536	block3d_dwcon			
v[0][0]				
block3d_activation (Activation) (None, 16, 16, 384) 0	block3d_bn[0]			
[0]				
block3d_se_squeeze (GlobalAvera (None, 384) 0	block3d_activ			
ation[0][0]				
block3d_se_reshape (Reshape) (None, 1, 1, 384) 0	block3d_se_sq			
ueeze[0][0]				
block3d_se_reduce (Conv2D) (None, 1, 1, 16) 6160	block3d_se_re			
shape[0][0]				
block3d_se_expand (Conv2D) (None, 1, 1, 384) 6528	block3d_se_re			
duce[0][0]				
block3d_se_excite (Multiply) (None, 16, 16, 384) 0	block3d_activ			
ation[0][0]				
block3d_se_ex pand[0][0]	block3d_se_ex			
block3d_project_conv (Conv2D) (None, 16, 16, 64) 24576	block3d_se_ex			

cate[0][0]

block3d_project_bn (BatchNormal (None, 16, 16, 64) 256	block3d_project_bn[0][0]	
block3d_drop (Dropout) (None, 16, 16, 64) 0	block3d_drop[0][0]	block3d_project_bn[0][0]
block3d_add (Add) (None, 16, 16, 64) 0	block3d_drop[0][0]	block3c_add[0][0]
block3e_expand_conv (Conv2D) (None, 16, 16, 384) 24576	block3d_add[0][0]	block3e_expand_conv[0][0]
block3e_expand_bn (BatchNormali (None, 16, 16, 384) 1536	block3e_expand_conv[0][0]	block3e_expand_conv[0][0]
block3e_expand_activation (Acti (None, 16, 16, 384) 0	block3e_expand_bn[0][0]	block3e_expand_activation[0][0]
block3e_dwconv (DepthwiseConv2D (None, 16, 16, 384) 9600	block3e_expand_activation[0][0]	block3e_dwconv[0][0]
block3e_bn (BatchNormalization) (None, 16, 16, 384) 1536	block3e_dwconv[0][0]	block3e_bn[0][0]
block3e_activation (Activation) (None, 16, 16, 384) 0	block3e_bn[0][0]	block3e_activation[0][0]
block3e_se_squeeze (GlobalAvera (None, 384) 0	block3e_activation[0][0]	block3e_activation[0][0]
block3e_se_reshape (Reshape) (None, 1, 1, 384) 0	block3e_se_squeeze[0][0]	block3e_se_squeeze[0][0]
block3e_se_reduce (Conv2D) (None, 1, 1, 16) 6160	block3e_se_reshape[0][0]	block3e_se_reduce[0][0]
block3e_se_expand (Conv2D) (None, 1, 1, 384) 6528	block3e_se_reduce[0][0]	block3e_se_reduce[0][0]
block3e_se_excite (Multiply) (None, 16, 16, 384) 0	block3e_se_reduce[0][0]	block3e_activation[0][0]
block3e_se_excite (Multiply) (None, 16, 16, 384) 0	block3e_activation[0][0]	block3e_se_excite[0][0]
block3e_project_conv (Conv2D) (None, 16, 16, 64) 24576	block3e_se_excite[0][0]	block3e_se_excite[0][0]
block3e_project_bn (BatchNormal (None, 16, 16, 64) 256	block3e_se_excite[0][0]	block3e_project_bn[0][0]

block3e_drop (Dropout) ct_bn[0][0]	(None, 16, 16, 64)	0	block3e_proje
block3e_add (Add) [0][0]	(None, 16, 16, 64)	0	block3e_drop block3d_add [0][0]
block4a_expand_conv (Conv2D) [0][0]	(None, 16, 16, 384)	24576	block3e_add
block4a_expand_bn (BatchNormali d_conv[0][0]	(None, 16, 16, 384)	1536	block4a_expan
block4a_expand_activation (Acti d_bn[0][0]	(None, 16, 16, 384)	0	block4a_expan
block4a_dwconv_pad (ZeroPadding _activation[0][0]	(None, 17, 17, 384)	0	block4a_expan
block4a_dwconv (DepthwiseConv2D v_pad[0][0]	(None, 8, 8, 384)	3456	block4a_dwcon
block4a_bn (BatchNormalizatio n)[0][0]	(None, 8, 8, 384)	1536	block4a_dwcon
block4a_activation (Activation) [0]	(None, 8, 8, 384)	0	block4a_bn[0]
block4a_se_squeeze (GlobalAvera tion[0][0]	(None, 384)	0	block4a_activ
block4a_se_reshape (Reshape) [0][0]	(None, 1, 1, 384)	0	block4a_se_sq ueeze[0][0]
block4a_se_reduce (Conv2D) shape[0][0]	(None, 1, 1, 16)	6160	block4a_se_re shape[0][0]
block4a_se_expand (Conv2D) duce[0][0]	(None, 1, 1, 384)	6528	block4a_se_re duce[0][0]
block4a_se_excite (Multiply) activation[0][0]	(None, 8, 8, 384)	0	block4a_activ ation[0][0]
block4a_project_conv (Conv2D) cite[0][0]	(None, 8, 8, 128)	49152	block4a_se_ex cite[0][0]
block4a_project_bn (BatchNormal ct_conv[0][0]	(None, 8, 8, 128)	512	block4a_proje

block4b_expand_conv (Conv2D) (None, 8, 8, 768)	98304	block4a_project_bn[0][0]
block4b_expand_bn (BatchNormali (None, 8, 8, 768)	3072	block4b_expand_conv[0][0]
block4b_expand_activation (Acti (None, 8, 8, 768)	0	block4b_expand_bn[0][0]
block4b_dwconv (DepthwiseConv2D (None, 8, 8, 768)	6912	block4b_expand_activation[0][0]
block4b_bn (BatchNormalizat (None, 8, 8, 768)	3072	block4b_dwconv[0][0]
block4b_activation (Activation) (None, 8, 8, 768)	0	block4b_bn[0][0]
block4b_se_squeeze (GlobalAvera (None, 768)	0	block4b_activation[0][0]
block4b_se_reshape (Reshape) (None, 1, 1, 768)	0	block4b_se_sq
block4b_se_reduce (Conv2D) (None, 1, 1, 32)	24608	ueeze[0][0] block4b_se_re
shape[0][0]		
block4b_se_expand (Conv2D) (None, 1, 1, 768)	25344	block4b_se_re
duce[0][0]		
block4b_se_excite (Multiply) (None, 8, 8, 768)	0	block4b_activ
ation[0][0]		
block4b_se_ex pand[0][0]		
block4b_project_conv (Conv2D) (None, 8, 8, 128)	98304	block4b_se_ex
cite[0][0]		
block4b_project_bn (BatchNormal (None, 8, 8, 128)	512	block4b_proje
ct_conv[0][0]		
block4b_drop (Dropout) (None, 8, 8, 128)	0	ct_bn[0][0] block4b_proje
block4b_add (Add) (None, 8, 8, 128)	0	block4b_drop[0][0]
block4a_project_bn[0][0]		
block4c_expand_conv (Conv2D) (None, 8, 8, 768)	98304	block4b_add[0][0]

block4c_expand_bn (BatchNormali (None, 8, 8, 768) 3072	block4c_expan
block4c_expand_activation (Acti (None, 8, 8, 768) 0	block4c_expan
block4c_dwconv (DepthwiseConv2D (None, 8, 8, 768) 6912	block4c_expan
block4c_bn (BatchNormalization) (None, 8, 8, 768) 3072	block4c_dwcon
block4c_activation (Activation) (None, 8, 8, 768) 0	block4c_bn[0]
block4c_se_squeeze (GlobalAvera (None, 768) 0	block4c_activ
block4c_se_reshape (Reshape) (None, 1, 1, 768) 0	block4c_se_sq
block4c_se_reduce (Conv2D) (None, 1, 1, 32) 24608	block4c_se_re
block4c_se_expand (Conv2D) (None, 1, 1, 768) 25344	block4c_se_re
block4c_se_excite (Multiply) (None, 8, 8, 768) 0	block4c_activ
block4c_se_ex (block4c_se_ex	pand[0][0]
block4c_project_conv (Conv2D) (None, 8, 8, 128) 98304	block4c_se_ex
block4c_project_bn (BatchNormal (None, 8, 8, 128) 512	block4c_proje
ct_conv[0][0]	
block4c_drop (Dropout) (None, 8, 8, 128) 0	block4c_proje
ct_bn[0][0]	
block4c_add (Add) (None, 8, 8, 128) 0	block4c_drop
[0][0]	block4b_add
block4d_expand_conv (Conv2D) (None, 8, 8, 768) 98304	block4c_add
[0][0]	
block4d_expand_bn (BatchNormali (None, 8, 8, 768) 3072	block4d_expan
block4d_conv[0][0]	
block4d_expand_activation (Acti (None, 8, 8, 768) 0	block4d_expan

d_bn[0][0]

block4d_dwconv (DepthwiseConv2D (None, 8, 8, 768)	6912	block4d_expand
d_activation[0][0]		
block4d_bn (BatchNormalization) (None, 8, 8, 768)	3072	block4d_dwconv
v[0][0]		
block4d_activation (Activation) (None, 8, 8, 768)	0	block4d_bn[0]
[0]		
block4d_se_squeeze (GlobalAvera (None, 768)	0	block4d_activ
ation[0][0]		
block4d_se_reshape (Reshape) (None, 1, 1, 768)	0	block4d_se_sq
ueeze[0][0]		
block4d_se_reduce (Conv2D) (None, 1, 1, 32)	24608	block4d_se_re
shape[0][0]		
block4d_se_expand (Conv2D) (None, 1, 1, 768)	25344	block4d_se_re
duce[0][0]		
block4d_se_excite (Multiply) (None, 8, 8, 768)	0	block4d_activ
ation[0][0]		
block4d_se_ex		
and[0][0]		
block4d_project_conv (Conv2D) (None, 8, 8, 128)	98304	block4d_se_ex
cite[0][0]		
block4d_project_bn (BatchNormal (None, 8, 8, 128)	512	block4d_proje
ct_conv[0][0]		
block4d_drop (Dropout) (None, 8, 8, 128)	0	block4d_proje
ct_bn[0][0]		
block4d_add (Add) (None, 8, 8, 128)	0	block4d_drop
[0][0]		
block4c_add		
[0][0]		
block4e_expand_conv (Conv2D) (None, 8, 8, 768)	98304	block4d_add
[0][0]		
block4e_expand_bn (BatchNormali (None, 8, 8, 768)	3072	block4e_expan
d_conv[0][0]		
block4e_expand_activation (Acti (None, 8, 8, 768)	0	block4e_expan
d_bn[0][0]		
block4e_dwconv (DepthwiseConv2D (None, 8, 8, 768)	6912	block4e_expan
d_activation[0][0]		

block4e_bn (BatchNormalization) (None, 8, 8, 768) v[0][0]	3072	block4e_dwcon
block4e_activation (Activation) (None, 8, 8, 768) [0]	0	block4e_bn[0]
block4e_se_squeeze (GlobalAvera (None, 768) activation[0][0]	0	block4e_activation
block4e_se_reshape (Reshape) (None, 1, 1, 768) ueeze[0][0]	0	block4e_se_sq
block4e_se_reduce (Conv2D) (None, 1, 1, 32) se_reshape[0][0]	24608	block4e_se_re
block4e_se_expand (Conv2D) (None, 1, 1, 768) duce[0][0]	25344	block4e_se_re
block4e_se_excite (Multiply) (None, 8, 8, 768) activation[0][0]	0	block4e_activ
block4e_se_ex pand[0][0]		block4e_se_ex
block4e_project_conv (Conv2D) (None, 8, 8, 128) project_conv[0][0]	98304	block4e_se_ex
block4e_project_bn (BatchNormal (None, 8, 8, 128) ct_conv[0][0]	512	block4e_proje
block4e_drop (Dropout) (None, 8, 8, 128) ct_bn[0][0]	0	block4e_proje
block4e_add (Add) [0][0]	(None, 8, 8, 128)	0
block4e_drop		block4e_drop
block4d_add		block4d_add
block4f_expand_conv (Conv2D) [0][0]	(None, 8, 8, 768)	98304
block4f_expand_bn (BatchNormali (None, 8, 8, 768) d_conv[0][0]	3072	block4f_expan
block4f_expand_activation (Acti (None, 8, 8, 768) d_bn[0][0]	0	block4f_expan
block4f_dwconv (DepthwiseConv2D (None, 8, 8, 768) d_activation[0][0]	6912	block4f_expan
block4f_bn (BatchNormalizat (None, 8, 8, 768) v[0][0]	3072	block4f_dwcon

block4f_activation (Activation) (None, 8, 8, 768) [0]	0	block4f_bn[0]
block4f_se_squeeze (GlobalAvera (None, 768) activation[0][0])	0	block4f_activation[0][0]
block4f_se_reshape (Reshape) (None, 1, 1, 768) reshape[0][0]	0	block4f_se_sq
block4f_se_reduce (Conv2D) (None, 1, 1, 32) reshape[0][0]	24608	block4f_se_re
block4f_se_expand (Conv2D) (None, 1, 1, 768) reduce[0][0]	25344	block4f_se_re
block4f_se_excite (Multiply) (None, 8, 8, 768) activation[0][0]	0	block4f_activ
		block4f_se_ex
block4f_project_conv (Conv2D) (None, 8, 8, 128) excite[0][0]	98304	block4f_se_ex
block4f_project_bn (BatchNormal (None, 8, 8, 128) conv[0][0])	512	block4f_proje
block4f_drop (Dropout) (None, 8, 8, 128) ct_bn[0][0]	0	block4f_proje
block4f_add (Add) (None, 8, 8, 128) [0][0]	0	block4f_drop
		block4e_add
block4g_expand_conv (Conv2D) (None, 8, 8, 768) [0][0]	98304	block4f_add
block4g_expand_bn (BatchNormali (None, 8, 8, 768) d_conv[0][0])	3072	block4g_expan
block4g_expand_activation (Acti (None, 8, 8, 768) d_bn[0][0])	0	block4g_expan
block4g_dwconv (DepthwiseConv2D (None, 8, 8, 768) d_activation[0][0])	6912	block4g_expan
block4g_bn (BatchNormalizat (None, 8, 8, 768) v[0][0])	3072	block4g_dwcon
block4g_activation (Activation) (None, 8, 8, 768) [0]	0	block4g_bn[0]

block4g_se_squeeze (GlobalAvera (None, 768) activation[0][0])	0	block4g_activation
block4g_se_reshape (Reshape) (None, 1, 1, 768)	0	block4g_se_sq
block4g_se_reduce (Conv2D) (None, 1, 1, 32)	24608	block4g_se_re
block4g_se_expand (Conv2D) (None, 1, 1, 768)	25344	block4g_se_re
block4g_se_excite (Multiply) (None, 8, 8, 768)	0	block4g_activation
block4g_se_ex		block4g_se_ex
block4g_project_conv (Conv2D) (None, 8, 8, 128)	98304	block4g_se_ex
block4g_project_bn (BatchNormal (None, 8, 8, 128)	512	block4g_proje
ct_conv[0][0])		ct
block4g_drop (Dropout) (None, 8, 8, 128)	0	block4g_proje
ct_bn[0][0])		ct
block4g_add (Add) (None, 8, 8, 128)	0	block4g_drop
[0][0])		block4f_add
block5a_expand_conv (Conv2D) (None, 8, 8, 768)	98304	block4g_add
[0][0])		[0]
block5a_expand_bn (BatchNormali (None, 8, 8, 768)	3072	block5a_expan
d_conv[0][0])		d
block5a_expand_activation (Acti (None, 8, 8, 768)	0	block5a_expan
d_bn[0][0])		d
block5a_dwconv (DepthwiseConv2D (None, 8, 8, 768)	19200	block5a_expan
activation[0][0])		activation
block5a_bn (BatchNormalization) (None, 8, 8, 768)	3072	block5a_dwcon
v[0][0])		v
block5a_activation (Activation) (None, 8, 8, 768)	0	block5a_bn[0]
[0])		[0]
block5a_se_squeeze (GlobalAvera (None, 768) activation[0][0])	0	block5a_activ
block5a_se_reshape (Reshape) (None, 1, 1, 768)	0	block5a_se_sq

ueeze[0][0]

block5a_se_reduce (Conv2D) shape[0][0]	(None, 1, 1, 32)	24608	block5a_se_re
block5a_se_expand (Conv2D) duce[0][0]	(None, 1, 1, 768)	25344	block5a_se_re
block5a_se_excite (Multiply) activation[0][0]	(None, 8, 8, 768)	0	block5a_activation block5a_se_ex
block5a_project_conv (Conv2D) cite[0][0]	(None, 8, 8, 176)	135168	block5a_se_ex
block5a_project_bn (BatchNormal ct_conv[0][0]	(None, 8, 8, 176)	704	block5a_proje
block5b_expand_conv (Conv2D) ct_bn[0][0]	(None, 8, 8, 1056)	185856	block5a_proje
block5b_expand_bn (BatchNormali d_conv[0][0]	(None, 8, 8, 1056)	4224	block5b_expan
block5b_expand_activation (Acti d_bn[0][0]	(None, 8, 8, 1056)	0	block5b_expan
block5b_dwconv (DepthwiseConv2D d_activation[0][0]	(None, 8, 8, 1056)	26400	block5b_expan
block5b_bn (BatchNormalizatio v[0][0]	(None, 8, 8, 1056)	4224	block5b_dwcon
block5b_activation (Activation) [0]	(None, 8, 8, 1056)	0	block5b_bn[0]
block5b_se_squeeze (GlobalAvera tion[0][0]	(None, 1056)	0	block5b_activ
block5b_se_reshape (Reshape) ueeze[0][0]	(None, 1, 1, 1056)	0	block5b_se_sq
block5b_se_reduce (Conv2D) shape[0][0]	(None, 1, 1, 44)	46508	block5b_se_re
block5b_se_expand (Conv2D) duce[0][0]	(None, 1, 1, 1056)	47520	block5b_se_re
block5b_se_excite (Multiply) activation[0][0]	(None, 8, 8, 1056)	0	block5b_activ block5b_se_ex
block5b_se_ex pand[0][0]			

block5b_project_conv (Conv2D) (None, 8, 8, 176)	185856	block5b_se_ex cite[0][0]
block5b_project_bn (BatchNormal (None, 8, 8, 176)	704	block5b_proje ct_conv[0][0]
block5b_drop (Dropout) (None, 8, 8, 176)	0	block5b_proje ct_bn[0][0]
block5b_add (Add) (None, 8, 8, 176)	0	block5b_drop block5a_proje ct_bn[0][0]
block5c_expand_conv (Conv2D) (None, 8, 8, 1056)	185856	block5b_add [0][0]
block5c_expand_bn (BatchNormali (None, 8, 8, 1056)	4224	block5c_expan d_conv[0][0]
block5c_expand_activation (Acti (None, 8, 8, 1056)	0	block5c_expan d_bn[0][0]
block5c_dwconv (DepthwiseConv2D (None, 8, 8, 1056)	26400	block5c_expan d_activation[0][0]
block5c_bn (BatchNormalization) (None, 8, 8, 1056)	4224	block5c_dwcon v[0][0]
block5c_activation (Activation) (None, 8, 8, 1056)	0	block5c_bn[0] [0]
block5c_se_squeeze (GlobalAvera (None, 1056)	0	block5c_activ ation[0][0]
block5c_se_reshape (Reshape) (None, 1, 1, 1056)	0	block5c_se_sq ueeze[0][0]
block5c_se_reduce (Conv2D) (None, 1, 1, 44)	46508	block5c_se_re shape[0][0]
block5c_se_expand (Conv2D) (None, 1, 1, 1056)	47520	block5c_se_re duce[0][0]
block5c_se_excite (Multiply) (None, 8, 8, 1056)	0	block5c_activ ation[0][0]
block5c_se_ex pand[0][0]		block5c_se_ex pand[0][0]
block5c_project_conv (Conv2D) (None, 8, 8, 176)	185856	block5c_se_ex cite[0][0]

block5c_project_bn (BatchNormal (None, 8, 8, 176)	704	block5c_proje
ct_conv[0][0]		
block5c_drop (Dropout) (None, 8, 8, 176)	0	block5c_proje
ct_bn[0][0]		
block5c_add (Add) (None, 8, 8, 176)	0	block5c_drop
[0][0]		block5b_add
block5c_add[0][0]		
block5d_expand_conv (Conv2D) (None, 8, 8, 1056)	185856	block5c_add
[0][0]		
block5d_expand_bn (BatchNormali (None, 8, 8, 1056)	4224	block5d_expan
conv[0][0]		d_conv[0][0]
block5d_expand_activation (Acti (None, 8, 8, 1056)	0	block5d_expan
ad_bn[0][0]		d_activation[0][0]
block5d_dwconv (DepthwiseConv2D (None, 8, 8, 1056)	26400	block5d_expan
activation[0][0]		d_activation[0][0]
block5d_bn (BatchNormalizat (None, 8, 8, 1056)	4224	block5d_dwcon
v[0][0]		v[0][0]
block5d_activation (Activation) (None, 8, 8, 1056)	0	block5d_bn[0]
[0]		
block5d_se_squeeze (GlobalAvera (None, 1056)	0	block5d_activ
ation[0][0]		ation[0][0]
block5d_se_reshape (Reshape) (None, 1, 1, 1056)	0	block5d_se_sq
ueeze[0][0]		ueeze[0][0]
block5d_se_reduce (Conv2D) (None, 1, 1, 44)	46508	block5d_se_re
shape[0][0]		shape[0][0]
block5d_se_expand (Conv2D) (None, 1, 1, 1056)	47520	block5d_se_re
duce[0][0]		duce[0][0]
block5d_se_excite (Multiply) (None, 8, 8, 1056)	0	block5d_activ
ation[0][0]		ation[0][0]
block5d_se_ex pand[0][0]		block5d_se_ex
pan[0][0]		pan[0][0]
block5d_project_conv (Conv2D) (None, 8, 8, 176)	185856	block5d_se_ex
cite[0][0]		cite[0][0]
block5d_project_bn (BatchNormal (None, 8, 8, 176)	704	block5d_proje
ct_conv[0][0]		ct_conv[0][0]

block5d_drop (Dropout) ct_bn[0][0]	(None, 8, 8, 176)	0	block5d_proje
block5d_add (Add) [0][0]	(None, 8, 8, 176)	0	block5d_drop block5c_add [0][0]
block5e_expand_conv (Conv2D) [0][0]	(None, 8, 8, 1056)	185856	block5d_add
block5e_expand_bn (BatchNormali d_conv[0][0]	(None, 8, 8, 1056)	4224	block5e_expan
block5e_expand_activation (Acti d_bn[0][0]	(None, 8, 8, 1056)	0	block5e_expan
block5e_dwconv (DepthwiseConv2D [0][0]	(None, 8, 8, 1056)	26400	block5e_expan
block5e_bn (BatchNormalizatio n)[0][0]	(None, 8, 8, 1056)	4224	block5e_dwcon
block5e_activation (Activation) [0]	(None, 8, 8, 1056)	0	block5e_bn[0]
block5e_se_squeeze (GlobalAvera ge)[0][0]	(None, 1, 1, 1056)	0	block5e_activ
block5e_se_reshape (Reshape) [0][0]	(None, 1, 1, 1056)	0	block5e_se_sq
block5e_se_reduce (Conv2D) shape[0][0]	(None, 1, 1, 44)	46508	block5e_se_re
block5e_se_expand (Conv2D) duce[0][0]	(None, 1, 1, 1056)	47520	block5e_se_re
block5e_se_excite (Multiply) ation[0][0]	(None, 8, 8, 1056)	0	block5e_activ
block5e_se_ex pand[0][0]			block5e_se_ex
block5e_project_conv (Conv2D) cite[0][0]	(None, 8, 8, 176)	185856	block5e_se_ex
block5e_project_bn (BatchNormal al_conv[0][0]	(None, 8, 8, 176)	704	block5e_proje
block5e_drop (Dropout) ct_bn[0][0]	(None, 8, 8, 176)	0	block5e_proje
block5e_add (Add)	(None, 8, 8, 176)	0	block5e_drop

[0][0]			block5d_add
[0][0]			
block5f_expand_conv (Conv2D)	(None, 8, 8, 1056)	185856	block5e_add
[0][0]			
block5f_expand_bn (BatchNormali	(None, 8, 8, 1056)	4224	block5f_expan
d_conv[0][0]			
block5f_expand_activation (Acti	(None, 8, 8, 1056)	0	block5f_expan
d_bn[0][0]			
block5f_dwconv (DepthwiseConv2D)	(None, 8, 8, 1056)	26400	block5f_expan
activation[0][0]			
block5f_bn (BatchNormalizatio	(None, 8, 8, 1056)	4224	block5f_dwcon
v[0][0]			
block5f_activation (Activation)	(None, 8, 8, 1056)	0	block5f_bn[0]
[0]			
block5f_se_squeeze (GlobalAvera	(None, 1056)	0	block5f_activ
activation[0][0]			
block5f_se_reshape (Reshape)	(None, 1, 1, 1056)	0	block5f_se_sq
ueeze[0][0]			
block5f_se_reduce (Conv2D)	(None, 1, 1, 44)	46508	block5f_se_re
shape[0][0]			
block5f_se_expand (Conv2D)	(None, 1, 1, 1056)	47520	block5f_se_re
duce[0][0]			
block5f_se_excite (Multiply)	(None, 8, 8, 1056)	0	block5f_activ
ation[0][0]			
block5f_se_ex			
pan[0][0]			
block5f_project_conv (Conv2D)	(None, 8, 8, 176)	185856	block5f_se_ex
cite[0][0]			
block5f_project_bn (BatchNormal	(None, 8, 8, 176)	704	block5f_proje
ct_conv[0][0]			
block5f_drop (Dropout)	(None, 8, 8, 176)	0	block5f_proje
ct_bn[0][0]			
block5f_add (Add)	(None, 8, 8, 176)	0	block5f_drop
[0][0]			
block5e_add			
[0][0]			

block5g_expand_conv (Conv2D) [0][0]	(None, 8, 8, 1056)	185856	block5f_add
block5g_expand_bn (BatchNormali [0][0]	(None, 8, 8, 1056)	4224	block5g_expan
block5g_expand_activation (Acti [0][0]	(None, 8, 8, 1056)	0	block5g_expan
block5g_dwconv (DepthwiseConv2D [0][0]	(None, 8, 8, 1056)	26400	block5g_expan
block5g_bn (BatchNormalizatio [0][0]	(None, 8, 8, 1056)	4224	block5g_dwcon
block5g_activation (Activation) [0]	(None, 8, 8, 1056)	0	block5g_bn[0]
block5g_se_squeeze (GlobalAvera [0][0]	(None, 1056)	0	block5g_activ
block5g_se_reshape (Reshape) [0][0]	(None, 1, 1, 1056)	0	block5g_se_sq
block5g_se_reduce (Conv2D) [0][0]	(None, 1, 1, 44)	46508	block5g_se_re
block5g_se_expand (Conv2D) [0][0]	(None, 1, 1, 1056)	47520	block5g_se_re
block5g_se_excite (Multiply) [0][0]	(None, 8, 8, 1056)	0	block5g_activ
block5g_se_ex pand[0][0]			block5g_se_ex
block5g_project_conv (Conv2D) [0][0]	(None, 8, 8, 176)	185856	block5g_se_ex
block5g_project_bn (BatchNormal [0][0]	(None, 8, 8, 176)	704	block5g_proje
block5g_drop (Dropout) [0][0]	(None, 8, 8, 176)	0	block5g_proje
block5g_add (Add) [0][0]	(None, 8, 8, 176)	0	block5g_drop
block5f_add [0][0]			block5f_add
block6a_expand_conv (Conv2D) [0][0]	(None, 8, 8, 1056)	185856	block5g_add
block6a_expand_bn (BatchNormali [0][0]	(None, 8, 8, 1056)	4224	block6a_expan

d_conv[0][0]

block6a_expand_activation (Activation) (None, 8, 8, 1056) 0 block6a_expand_bn[0][0]

block6a_dwconv_pad (ZeroPadding) (None, 11, 11, 1056) 0 block6a_expand_activation[0][0]

block6a_dwconv (DepthwiseConv2D) (None, 4, 4, 1056) 26400 block6a_dwconv_pad[0][0]

block6a_bn (BatchNormalization) (None, 4, 4, 1056) 4224 block6a_dwconv_v[0][0]

block6a_activation (Activation) (None, 4, 4, 1056) 0 block6a_bn[0][0]

block6a_se_squeeze (GlobalAverage) (None, 1056) 0 block6a_activation[0][0]

block6a_se_reshape (Reshape) (None, 1, 1, 1056) 0 block6a_se_sq

ueeze[0][0]

block6a_se_reduce (Conv2D) (None, 1, 1, 44) 46508 block6a_se_reshape[0][0]

block6a_se_expand (Conv2D) (None, 1, 1, 1056) 47520 block6a_se_re

duce[0][0]

block6a_se_excite (Multiply) (None, 4, 4, 1056) 0 block6a_activation[0][0]

block6a_se_ex

pand[0][0]

block6a_project_conv (Conv2D) (None, 4, 4, 304) 321024 block6a_se_ex

cite[0][0]

block6a_project_bn (BatchNormal) (None, 4, 4, 304) 1216 block6a_pro

ject_conv[0][0]

block6b_expand_conv (Conv2D) (None, 4, 4, 1824) 554496 block6a_pro

ject_bn[0][0]

block6b_expand_bn (BatchNormali (None, 4, 4, 1824) 7296 block6b_expan

d_conv[0][0]

block6b_expand_activation (Activation) (None, 4, 4, 1824) 0 block6b_expan

d_bn[0][0]

block6b_dwconv (DepthwiseConv2D) (None, 4, 4, 1824) 45600 block6b_expan

d_activation[0][0]

block6b_bn (BatchNormalization) (None, 4, 4, 1824) v[0][0]	7296	block6b_dwcon
block6b_activation (Activation) (None, 4, 4, 1824) [0]	0	block6b_bn[0]
block6b_se_squeeze (GlobalAvera (None, 1824) activation[0][0]	0	block6b_activ
block6b_se_reshape (Reshape) (None, 1, 1, 1824) squeeze[0][0]	0	block6b_se_sq
block6b_se_reduce (Conv2D) (None, 1, 1, 76) se_reduce[0][0]	138700	block6b_se_re
block6b_se_expand (Conv2D) (None, 1, 1, 1824) se_reduce[0][0]	140448	block6b_se_re
block6b_se_excite (Multiply) (None, 4, 4, 1824) activation[0][0]	0	block6b_activ
block6b_se_excite (Multiply) (None, 4, 4, 1824) expand[0][0]	0	block6b_se_ex
block6b_project_conv (Conv2D) (None, 4, 4, 304) project_conv[0][0]	554496	block6b_se_ex
block6b_project_bn (BatchNormal (None, 4, 4, 304) project_conv[0][0]	1216	block6b_proje
block6b_drop (Dropout) (None, 4, 4, 304) project_bn[0][0]	0	block6b_proje
block6b_add (Add) [0][0]	(None, 4, 4, 304)	0
block6b_add (Add) [0][0]	(None, 4, 4, 304)	0
block6b_add (Add) [0][0]	(None, 4, 4, 304)	0
block6c_expand_conv (Conv2D) (None, 4, 4, 1824) [0][0]	554496	block6b_add
block6c_expand_bn (BatchNormali (None, 4, 4, 1824) expand_conv[0][0]	7296	block6c_expan
block6c_expand_activation (Acti (None, 4, 4, 1824) d_bn[0][0]	0	block6c_expan
block6c_dwconv (DepthwiseConv2D (None, 4, 4, 1824) d_activation[0][0]	45600	block6c_expan
block6c_bn (BatchNormaliz (None, 4, 4, 1824) v[0][0]	7296	block6c_dwcon
block6c_activation (Activation) (None, 4, 4, 1824) [0]	0	block6c_bn[0]

[0]

block6c_se_squeeze (GlobalAvera (None, 1824)	0	block6c_activation[0][0]
block6c_se_reshape (Reshape) (None, 1, 1, 1824)	0	block6c_se_sq ueeze[0][0]
block6c_se_reduce (Conv2D) (None, 1, 1, 76)	138700	block6c_se_re shape[0][0]
block6c_se_expand (Conv2D) (None, 1, 1, 1824)	140448	block6c_se_re duce[0][0]
block6c_se_excite (Multiply) (None, 4, 4, 1824)	0	block6c_activ ation[0][0]
block6c_se_ex pand[0][0]		
block6c_project_conv (Conv2D) (None, 4, 4, 304)	554496	block6c_se_ex cite[0][0]
block6c_project_bn (BatchNormal (None, 4, 4, 304)	1216	block6c_proje ct_conv[0][0]
block6c_drop (Dropout) (None, 4, 4, 304)	0	block6c_proje ct_bn[0][0]
block6c_add (Add) [0][0]	0	block6c_drop [0][0]
block6c_add		block6b_add
block6d_expand_conv (Conv2D) (None, 4, 4, 1824)	554496	block6c_add [0][0]
block6d_expand_bn (BatchNormali (None, 4, 4, 1824)	7296	block6d_expan d_conv[0][0]
block6d_expand_activation (Acti (None, 4, 4, 1824)	0	block6d_expan d_bn[0][0]
block6d_dwconv (DepthwiseConv2D (None, 4, 4, 1824)	45600	block6d_expan d_activation[0][0]
block6d_bn (BatchNormalizat (None, 4, 4, 1824)	7296	block6d_dwcon v[0][0]
block6d_activation (Activation) (None, 4, 4, 1824)	0	block6d_bn[0] [0]
block6d_se_squeeze (GlobalAvera (None, 1824)	0	block6d_activ ation[0][0]

block6d_se_reshape (Reshape) <code>zeeze[0][0]</code>	(None, 1, 1, 1824)	0	block6d_se_sq
block6d_se_reduce (Conv2D) <code>shape[0][0]</code>	(None, 1, 1, 76)	138700	block6d_se_re
block6d_se_expand (Conv2D) <code>duce[0][0]</code>	(None, 1, 1, 1824)	140448	block6d_se_re
block6d_se_excite (Multiply) <code>activation[0][0]</code>	(None, 4, 4, 1824)	0	block6d_activ
block6d_se_excite[0][0]			block6d_se_ex
block6d_project_conv (Conv2D) <code>cite[0][0]</code>	(None, 4, 4, 304)	554496	block6d_se_ex
block6d_project_bn (BatchNormal <code>ct_conv[0][0]</code>	(None, 4, 4, 304)	1216	block6d_proje
block6d_drop (Dropout) <code>ct_bn[0][0]</code>	(None, 4, 4, 304)	0	block6d_proje
block6d_add (Add) <code>[0][0]</code>	(None, 4, 4, 304)	0	block6d_drop
block6c_add <code>[0][0]</code>			block6c_add
block6e_expand_conv (Conv2D) <code>[0][0]</code>	(None, 4, 4, 1824)	554496	block6d_add
block6e_expand_bn (BatchNormali <code>d_conv[0][0]</code>	(None, 4, 4, 1824)	7296	block6e_expan
block6e_expand_activation (Acti <code>d_bn[0][0]</code>	(None, 4, 4, 1824)	0	block6e_expan
block6e_dwconv (DepthwiseConv2D) <code>activation[0][0]</code>	(None, 4, 4, 1824)	45600	block6e_expan
block6e_bn (BatchNormalizatio <code>v[0][0]</code>	(None, 4, 4, 1824)	7296	block6e_dwcon
block6e_activation (Activation) <code>[0]</code>	(None, 4, 4, 1824)	0	block6e_bn[0]
block6e_se_squeeze (GlobalAvera <code>ation[0][0]</code>	(None, 1824)	0	block6e_activ
block6e_se_reshape (Reshape) <code>ueueze[0][0]</code>	(None, 1, 1, 1824)	0	block6e_se_sq

block6e_se_reduce (Conv2D) shape[0][0]	(None, 1, 1, 76)	138700	block6e_se_re
block6e_se_expand (Conv2D) duce[0][0]	(None, 1, 1, 1824)	140448	block6e_se_re
block6e_se_excite (Multiply) activation[0][0]	(None, 4, 4, 1824)	0	block6e_activation block6e_se_ex
block6e_project_conv (Conv2D) cite[0][0]	(None, 4, 4, 304)	554496	block6e_se_ex
block6e_project_bn (BatchNormal ct_conv[0][0]	(None, 4, 4, 304)	1216	block6e_proje
block6e_drop (Dropout) ct_bn[0][0]	(None, 4, 4, 304)	0	block6e_proje
block6e_add (Add) [0][0]	(None, 4, 4, 304)	0	block6e_drop block6d_add
block6f_expand_conv (Conv2D) [0][0]	(None, 4, 4, 1824)	554496	block6e_add
block6f_expand_bn (BatchNormali d_conv[0][0]	(None, 4, 4, 1824)	7296	block6f_expan
block6f_expand_activation (Acti d_bn[0][0]	(None, 4, 4, 1824)	0	block6f_expan
block6f_dwconv (DepthwiseConv2D) activation[0][0]	(None, 4, 4, 1824)	45600	block6f_expan
block6f_bn (BatchNormalizatio v[0][0]	(None, 4, 4, 1824)	7296	block6f_dwcon
block6f_activation (Activation) [0]	(None, 4, 4, 1824)	0	block6f_bn[0]
block6f_se_squeeze (GlobalAvera activation[0][0]	(None, 1824)	0	block6f_activ
block6f_se_reshape (Reshape) squeez[0][0]	(None, 1, 1, 1824)	0	block6f_se_sq
block6f_se_reduce (Conv2D) shape[0][0]	(None, 1, 1, 76)	138700	block6f_se_re

block6f_se_expand (Conv2D) duce[0][0]	(None, 1, 1, 1824)	140448	block6f_se_re
block6f_se_excite (Multiply) ation[0][0]	(None, 4, 4, 1824)	0	block6f_activ block6f_se_ex
block6f_project_conv (Conv2D) cite[0][0]	(None, 4, 4, 304)	554496	block6f_se_ex
block6f_project_bn (BatchNormal ct_conv[0][0]	(None, 4, 4, 304)	1216	block6f_proje
block6f_drop (Dropout) ct_bn[0][0]	(None, 4, 4, 304)	0	block6f_proje
block6f_add (Add) [0][0]	(None, 4, 4, 304)	0	block6f_drop block6e_add
block6g_expand_conv (Conv2D) [0][0]	(None, 4, 4, 1824)	554496	block6f_add
block6g_expand_bn (BatchNormali d_conv[0][0]	(None, 4, 4, 1824)	7296	block6g_expan
block6g_expand_activation (Acti d_bn[0][0]	(None, 4, 4, 1824)	0	block6g_expan
block6g_dwconv (DepthwiseConv2D) activation[0][0]	(None, 4, 4, 1824)	45600	block6g_expan
block6g_bn (BatchNormalizatio v[0][0]	(None, 4, 4, 1824)	7296	block6g_dwcon
block6g_activation (Activation) [0]	(None, 4, 4, 1824)	0	block6g_bn[0]
block6g_se_squeeze (GlobalAvera tion[0][0]	(None, 1824)	0	block6g_activ
block6g_se_reshape (Reshape) squeez[0][0]	(None, 1, 1, 1824)	0	block6g_se_sq
block6g_se_reduce (Conv2D) shape[0][0]	(None, 1, 1, 76)	138700	block6g_se_re
block6g_se_expand (Conv2D) duce[0][0]	(None, 1, 1, 1824)	140448	block6g_se_re
block6g_se_excite (Multiply)	(None, 4, 4, 1824)	0	block6g_activ

ation[0][0]				block6g_se_ex
pand[0][0]				
block6g_project_conv (Conv2D) (None, 4, 4, 304)	554496			block6g_se_ex
cite[0][0]				
block6g_project_bn (BatchNormal (None, 4, 4, 304)	1216			block6g_proje
ct_conv[0][0]				
block6g_drop (Dropout) (None, 4, 4, 304)	0			block6g_proje
ct_bn[0][0]				
block6g_add (Add) (None, 4, 4, 304)	0			block6g_drop
[0][0]				block6f_add
[0][0]				
block6h_expand_conv (Conv2D) (None, 4, 4, 1824)	554496			block6g_add
[0][0]				
block6h_expand_bn (BatchNormali (None, 4, 4, 1824)	7296			block6h_expan
d_conv[0][0]				
block6h_expand_activation (Acti (None, 4, 4, 1824)	0			block6h_expan
d_bn[0][0]				
block6h_dwconv (DepthwiseConv2D (None, 4, 4, 1824)	45600			block6h_expan
d_activation[0][0]				
block6h_bn (BatchNormalizati (None, 4, 4, 1824)	7296			block6h_dwcon
v[0][0]				
block6h_activation (Activation) (None, 4, 4, 1824)	0			block6h_bn[0]
[0]				
block6h_se_squeeze (GlobalAvera (None, 1824)	0			block6h_activ
ation[0][0]				
block6h_se_reshape (Reshape) (None, 1, 1, 1824)	0			block6h_se_sq
ueeze[0][0]				
block6h_se_reduce (Conv2D) (None, 1, 1, 76)	138700			block6h_se_re
shape[0][0]				
block6h_se_expand (Conv2D) (None, 1, 1, 1824)	140448			block6h_se_re
duce[0][0]				
block6h_se_excite (Multiply) (None, 4, 4, 1824)	0			block6h_activ
ation[0][0]				
pand[0][0]				block6h_se_ex

block6h_project_conv (Conv2D) (None, 4, 4, 304)	554496	block6h_se_ex
cite[0][0]		
block6h_project_bn (BatchNormal (None, 4, 4, 304)	1216	block6h_proje
ct_conv[0][0]		
block6h_drop (Dropout) (None, 4, 4, 304)	0	block6h_proje
ct_bn[0][0]		
block6h_add (Add) (None, 4, 4, 304)	0	block6h_drop
[0][0]		block6g_add
block6i_expand_conv (Conv2D) (None, 4, 4, 1824)	554496	block6h_add
[0][0]		
block6i_expand_bn (BatchNormali (None, 4, 4, 1824)	7296	block6i_expan
d_conv[0][0]		d_conv[0][0]
block6i_expand_activation (Acti (None, 4, 4, 1824)	0	block6i_expan
d_bn[0][0]		d_activation[0][0]
block6i_dwconv (DepthwiseConv2D (None, 4, 4, 1824)	45600	block6i_expan
d_activation[0][0]		d_activation[0][0]
block6i_bn (BatchNormalization) (None, 4, 4, 1824)	7296	block6i_dwcon
v[0][0]		v[0][0]
block6i_activation (Activation) (None, 4, 4, 1824)	0	block6i_bn[0]
[0]		
block6i_se_squeeze (GlobalAvera (None, 1824)	0	block6i_activ
tation[0][0]		ation[0][0]
block6i_se_reshape (Reshape) (None, 1, 1, 1824)	0	block6i_se_sq
ueeze[0][0]		ueeze[0][0]
block6i_se_reduce (Conv2D) (None, 1, 1, 76)	138700	block6i_se_re
shape[0][0]		shape[0][0]
block6i_se_expand (Conv2D) (None, 1, 1, 1824)	140448	block6i_se_re
duce[0][0]		duce[0][0]
block6i_se_excite (Multiply) (None, 4, 4, 1824)	0	block6i_activ
tation[0][0]		tation[0][0]
block6i_se_ex (block6i_se_ex)		
pand[0][0]		pand[0][0]
block6i_project_conv (Conv2D) (None, 4, 4, 304)	554496	block6i_se_ex
cite[0][0]		
block6i_project_bn (BatchNormal (None, 4, 4, 304)	1216	block6i_proje
ct[0][0]		

ct_conv[0][0]

block6i_drop (Dropout) ct_bn[0][0]	(None, 4, 4, 304)	0	block6i_proje
block6i_add (Add) [0][0]	(None, 4, 4, 304)	0	block6i_drop block6h_add [0][0]
block7a_expand_conv (Conv2D) [0][0]	(None, 4, 4, 1824)	554496	block6i_add
block7a_expand_bn (BatchNormali d_conv[0][0]	(None, 4, 4, 1824)	7296	block7a_expan
block7a_expand_activation (Acti d_bn[0][0]	(None, 4, 4, 1824)	0	block7a_expan
block7a_dwconv (DepthwiseConv2D activation[0][0]	(None, 4, 4, 1824)	16416	block7a_expan
block7a_bn (BatchNormalizatio v[0][0]	(None, 4, 4, 1824)	7296	block7a_dwcon
block7a_activation (Activation) [0]	(None, 4, 4, 1824)	0	block7a_bn[0]
block7a_se_squeeze (GlobalAvera tion[0][0]	(None, 1, 1, 1824)	0	block7a_activ
block7a_se_reshape (Reshape) [0][0]	(None, 1, 1, 1824)	0	block7a_se_sq
block7a_se_reduce (Conv2D) shape[0][0]	(None, 1, 1, 76)	138700	block7a_se_re
block7a_se_expand (Conv2D) duce[0][0]	(None, 1, 1, 1824)	140448	block7a_se_re
block7a_se_excite (Multiply) activation[0][0]	(None, 4, 4, 1824)	0	block7a_activ
block7a_se_ex pand[0][0]			block7a_se_ex
block7a_project_conv (Conv2D) cite[0][0]	(None, 4, 4, 512)	933888	block7a_se_ex
block7a_project_bn (BatchNormal ct_conv[0][0]	(None, 4, 4, 512)	2048	block7a_proje
block7b_expand_conv (Conv2D) ct_bn[0][0]	(None, 4, 4, 3072)	1572864	block7a_proje

block7b_expand_bn (BatchNormali (None, 4, 4, 3072)	12288	block7b_expand_conv[0][0]
block7b_expand_activation (Acti (None, 4, 4, 3072)	0	block7b_expand_bn[0][0]
block7b_dwconv (DepthwiseConv2D (None, 4, 4, 3072)	27648	block7b_expand_activation[0][0]
block7b_bn (BatchNormalization) (None, 4, 4, 3072)	12288	block7b_dwconv[0][0]
block7b_activation (Activation) (None, 4, 4, 3072)	0	block7b_bn[0][0]
block7b_se_squeeze (GlobalAvera (None, 3072)	0	block7b_activation[0][0]
block7b_se_reshape (Reshape) (None, 1, 1, 3072)	0	block7b_se_sq
block7b_se_reduce (Conv2D) (None, 1, 1, 128)	393344	block7b_se_re
block7b_se_expand (Conv2D) (None, 1, 1, 3072)	396288	shape[0][0] block7b_se_re
block7b_se_excite (Multiply) (None, 4, 4, 3072)	0	shape[0][0] block7b_activation[0][0]
block7b_se_ex pand[0][0]		
block7b_project_conv (Conv2D) (None, 4, 4, 512)	1572864	block7b_se_ex
block7b_project_bn (BatchNormal (None, 4, 4, 512)	2048	cite[0][0] block7b_proje
ct_conv[0][0]		
block7b_drop (Dropout) (None, 4, 4, 512)	0	ct_bn[0][0] block7b_proje
ct_bn[0][0]		
block7b_add (Add) (None, 4, 4, 512)	0	block7b_drop[0][0] block7a_proje
ct_bn[0][0]		
block7c_expand_conv (Conv2D) (None, 4, 4, 3072)	1572864	block7b_add[0][0]
block7c_expand_bn (BatchNormali (None, 4, 4, 3072)	12288	block7c_expand_conv[0][0]

block7c_expand_activation (Acti	(None, 4, 4, 3072)	0	block7c_expand
d_bn[0][0]			
block7c_dwconv (DepthwiseConv2D	(None, 4, 4, 3072)	27648	block7c_expand
activation[0][0]			
block7c_bn (BatchNormalizatio	(None, 4, 4, 3072)	12288	block7c_dwcon
v[0][0]			
block7c_activation (Activatio	(None, 4, 4, 3072)	0	block7c_bn[0]
[0]			
block7c_se_squeeze (GlobalAvera	(None, 3072)	0	block7c_activa
tion[0][0]			
block7c_se_reshape (Reshape)	(None, 1, 1, 3072)	0	block7c_se_sq
ueeze[0][0]			
block7c_se_reduce (Conv2D)	(None, 1, 1, 128)	393344	block7c_se_re
shape[0][0]			
block7c_se_expand (Conv2D)	(None, 1, 1, 3072)	396288	block7c_se_re
duce[0][0]			
block7c_se_excite (Multipli	(None, 4, 4, 3072)	0	block7c_activa
tion[0][0]			
block7c_se_ex			
pand[0][0]			
block7c_project_conv (Conv2D)	(None, 4, 4, 512)	1572864	block7c_se_ex
cite[0][0]			
block7c_project_bn (BatchNorma	(None, 4, 4, 512)	2048	block7c_proje
l_conv[0][0]			
block7c_drop (Dropout)	(None, 4, 4, 512)	0	block7c_proje
ct_bn[0][0]			
block7c_add (Add)	(None, 4, 4, 512)	0	block7c_drop
[0][0]			
block7b_add			
[0][0]			
top_conv (Conv2D)	(None, 4, 4, 2048)	1048576	block7c_add
[0][0]			
top_bn (BatchNormalizatio	(None, 4, 4, 2048)	8192	top_conv[0]
n[0]			
top_activation (Activatio	(None, 4, 4, 2048)	0	top_bn[0][0]
n[0]			
global_average_pooling2d (Globa	(None, 2048)	0	top_activatio
l[0]			

```
n[0][0]
```

batch_normalization_6 (BatchNor (None, 2048)	8192	global_averag
e_pooling2d[0][0]		
dropout_4 (Dropout) (None, 2048)	0	batch_normaliz
ation_6[0][0]		
dense_2 (Dense) (None, 32)	65568	dropout_4[0]
[0]		
dense_3 (Dense) (None, 1)	33	dense_2[0][0]
=====	=====	=====
Total params: 28,587,320		
Trainable params: 69,697		
Non-trainable params: 28,517,623		

```
In [22]:
```

```
model_2.compile(  
    loss = 'mse',  
    optimizer = 'Adam',  
    metrics = [tf.keras.metrics.RootMeanSquaredError(name="rmse")])
```

```
In [23]:
```

```
from tensorflow.keras.preprocessing.image import ImageDataGenerator  
data_augmentation = ImageDataGenerator(rotation_range=15, zoom_range=0.15, width_shift_range = 0.2,  
height_shift_range = 0.2,  
shear_range = 0.1,  
horizontal_flip = True,  
fill_mode = "nearest")
```

```
In [24]:
```

```
kall = keras.callbacks.EarlyStopping(monitor='val_rmse', patience=10, restore_best_weights=True)  
  
history = model_2.fit(  
    data_augmentation.flow(x_train,y_train,batch_size=1024),  
    validation_data = (x_test,y_test),  
    steps_per_epoch = len(x_train) // 1024,  
    epochs = 30, callbacks=[kall]  
)
```

```
Epoch 1/30  
6/6 [=====] - 432s 74s/step - loss: 1873.9289 - rmse: 43.2888 - val_loss: 1857.3164 - val_rmse: 43.0966  
Epoch 2/30  
6/6 [=====] - 410s 72s/step - loss: 1812.0355 - rmse: 42.5676 - val_loss: 1906.5901 - val_rmse: 43.6645  
Epoch 3/30  
6/6 [=====] - 400s 70s/step - loss: 1827.2462 - rmse: 42.7447 - val_loss: 1753.0638 - val_rmse: 41.8696  
Epoch 4/30  
6/6 [=====] - 475s 85s/step - loss: 1709.9631 - rmse: 41.3512 - val_loss: 1771.9995 - val_rmse: 42.0951  
Epoch 5/30  
6/6 [=====] - 469s 78s/step - loss: 1657.8936 - rmse: 40.7161 - val_loss: 1636.2598 - val_rmse: 40.4507  
Epoch 6/30  
6/6 [=====] - 396s 69s/step - loss: 1568.8812 - rmse: 39.6011 - val_loss: 1463.4620 - val_rmse: 38.2552  
Epoch 7/30
```

6/6 [=====] - 395s 69s/step - loss: 1375.3240 - rmse: 37.0799 - val_loss: 1272.5067 - val_rmse: 35.6722
Epoch 8/30
6/6 [=====] - 416s 76s/step - loss: 1146.2941 - rmse: 33.8539 - val_loss: 1021.0345 - val_rmse: 31.9536
Epoch 9/30
6/6 [=====] - 407s 71s/step - loss: 913.2666 - rmse: 30.2130 - val_loss: 734.3057 - val_rmse: 27.0981
Epoch 10/30
6/6 [=====] - 401s 70s/step - loss: 639.5902 - rmse: 25.2877 - val_loss: 527.7954 - val_rmse: 22.9738
Epoch 11/30
6/6 [=====] - 401s 70s/step - loss: 484.8349 - rmse: 22.0133 - val_loss: 437.0920 - val_rmse: 20.9067
Epoch 12/30
6/6 [=====] - 399s 70s/step - loss: 447.2597 - rmse: 21.1478 - val_loss: 432.1925 - val_rmse: 20.7892
Epoch 13/30
6/6 [=====] - 398s 70s/step - loss: 445.9243 - rmse: 21.1168 - val_loss: 435.8708 - val_rmse: 20.8775
Epoch 14/30
6/6 [=====] - 425s 73s/step - loss: 467.3221 - rmse: 21.6163 - val_loss: 431.2438 - val_rmse: 20.7664
Epoch 15/30
6/6 [=====] - 400s 70s/step - loss: 435.7649 - rmse: 20.8749 - val_loss: 427.6853 - val_rmse: 20.6806
Epoch 16/30
6/6 [=====] - 400s 70s/step - loss: 424.7081 - rmse: 20.6065 - val_loss: 429.6132 - val_rmse: 20.7271
Epoch 17/30
6/6 [=====] - 401s 70s/step - loss: 434.2966 - rmse: 20.8385 - val_loss: 433.6300 - val_rmse: 20.8238
Epoch 18/30
6/6 [=====] - 399s 70s/step - loss: 455.0258 - rmse: 21.3283 - val_loss: 430.5573 - val_rmse: 20.7499
Epoch 19/30
6/6 [=====] - 400s 70s/step - loss: 446.2630 - rmse: 21.1245 - val_loss: 428.0292 - val_rmse: 20.6889
Epoch 20/30
6/6 [=====] - 399s 70s/step - loss: 449.3381 - rmse: 21.1955 - val_loss: 427.4152 - val_rmse: 20.6740
Epoch 21/30
6/6 [=====] - 399s 70s/step - loss: 446.8055 - rmse: 21.1369 - val_loss: 427.4100 - val_rmse: 20.6739
Epoch 22/30
6/6 [=====] - 441s 79s/step - loss: 448.6512 - rmse: 21.1781 - val_loss: 428.5179 - val_rmse: 20.7007
Epoch 23/30
6/6 [=====] - 440s 77s/step - loss: 463.0178 - rmse: 21.5127 - val_loss: 428.7200 - val_rmse: 20.7056
Epoch 24/30
6/6 [=====] - 412s 71s/step - loss: 442.0092 - rmse: 21.0209 - val_loss: 428.1041 - val_rmse: 20.6907
Epoch 25/30
6/6 [=====] - 400s 70s/step - loss: 429.3829 - rmse: 20.7176 - val_loss: 429.2787 - val_rmse: 20.7190
Epoch 26/30
6/6 [=====] - 407s 72s/step - loss: 449.7387 - rmse: 21.2047 - val_loss: 428.5323 - val_rmse: 20.7010
Epoch 27/30
6/6 [=====] - 403s 70s/step - loss: 435.2628 - rmse: 20.8623 - val_loss: 427.5097 - val_rmse: 20.6763
Epoch 28/30
6/6 [=====] - 397s 70s/step - loss: 435.1812 - rmse: 20.8608 - val_loss: 427.4426 - val_rmse: 20.6747
Epoch 29/30
6/6 [=====] - 395s 72s/step - loss: 428.4867 - rmse: 20.6995 - val_loss: 427.9561 - val_rmse: 20.6871
Epoch 30/30

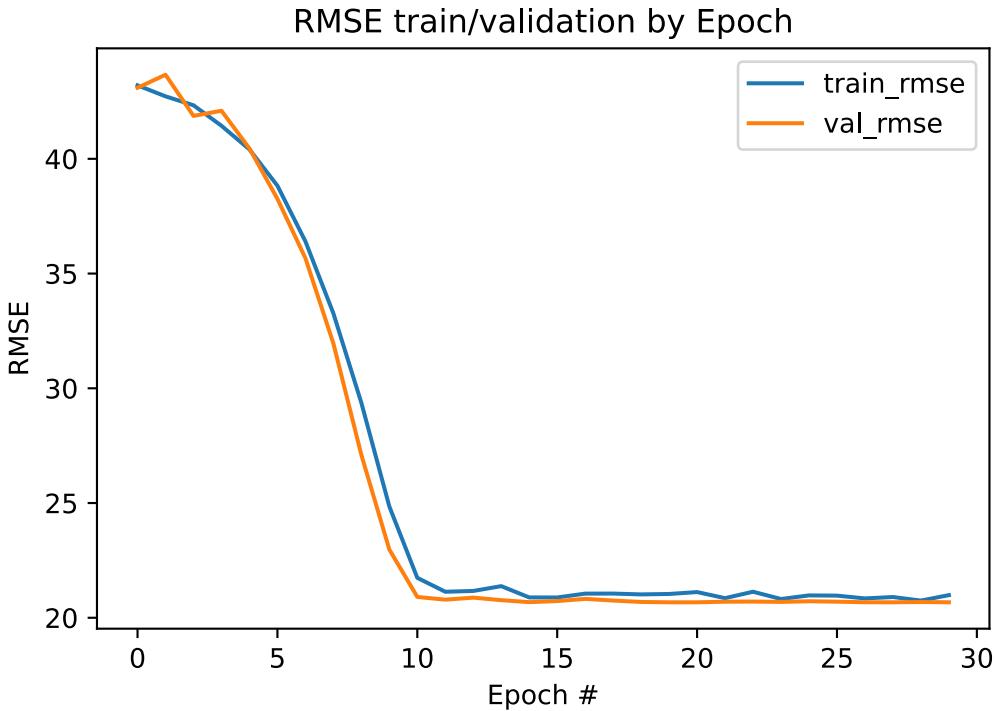
```
6/6 [=====] - 419s 74s/step - loss: 448.2925 - rmse: 21.1716 - val_loss: 427.4606 - val_rmse: 20.6751
```

We got a rmse value of around 20.6 after 30 epochs, which is slightly better than model1

In [25]:

```
plt.figure()
plt.plot(history.history["rmse"], label="train_rmse")
plt.plot(history.history["val_rmse"], label="val_rmse")
plt.title("RMSE train/validation by Epoch")
plt.xlabel("Epoch #")
plt.ylabel("RMSE")
plt.legend(loc="upper right")
```

Out[25]: <matplotlib.legend.Legend at 0x1624448f208>



In [26]:

```
model_2.save('my_model2')
```

```
INFO:tensorflow:Assets written to: my_model2\assets
```

```
INFO:tensorflow:Assets written to: my_model2\assets
```

In [87]:

```
reload_model_2 = tf.keras.models.load_model('my_model2')
```

```
WARNING:absl:Importing a function (_inference_block4g_se_reduce_layer_call_and_return_conditional_losses_104385) with ops with custom gradients. Will likely fail if a gradient is requested.
```

```
WARNING:absl:Importing a function (_inference_block2a_activation_layer_call_and_return_conditional_losses_97455) with ops with custom gradients. Will likely fail if a gradient is requested.
```

```
WARNING:absl:Importing a function (_inference_block4f_expand_activation_layer_call_and_return_conditional_losses_103806) with ops with custom gradients. Will likely fail if a gradient is requested.
```

```
WARNING:absl:Importing a function (_inference_block3c_expand_activation_layer_call_and_return_conditional_losses_140278) with ops with custom gradients. Will likely fail if a gradient is requested.
```

```
WARNING:absl:Importing a function (_inference_block5c_expand_activation_layer_call_and_return_conditional_losses_105513) with ops with custom gradients. Will likely fail if a gradient is requested.
```

```
WARNING:absl:Importing a function (_inference_block2d_se_reduce_layer_call_and_return_conditional_losses_138292) with ops with custom gradients. Will likely fail if a gradient is requested.
```


WARNING:absl:Importing a function (`__inference_block5d_activation_layer_call_and_return_conditional_losses_106045`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6d_se_reduce_layer_call_and_return_conditional_losses_109120`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block2c_se_reduce_layer_call_and_return_conditional_losses_98329`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block2a_se_reduce_layer_call_and_return_conditional_losses_136680`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6e_expand_activation_layer_call_and_return_conditional_losses_109421`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block4e_se_reduce_layer_call_and_return_conditional_losses_144281`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block7b_se_reduce_layer_call_and_return_conditional_losses_112147`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block4b_se_reduce_layer_call_and_return_conditional_losses_102185`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block4e_activation_layer_call_and_return_conditional_losses_103458`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block4g_activation_layer_call_and_return_conditional_losses_145347`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6g_expand_activation_layer_call_and_return_conditional_losses_110301`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6f_expand_activation_layer_call_and_return_conditional_losses_109861`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3e_activation_layer_call_and_return_conditional_losses_141523`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block4a_expand_activation_layer_call_and_return_conditional_losses_141937`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6c_se_reduce_layer_call_and_return_conditional_losses_108680`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3b_expand_activation_layer_call_and_return_conditional_losses_99898`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block2d_se_reduce_layer_call_and_return_conditional_losses_98769`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block5b_activation_layer_call_and_return_conditional_losses_146406`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3b_activation_layer_call_and_return_conditional_losses_99990`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block5e_activation_layer_call_and_return_conditional_losses_148065`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block4g_expand_activation_layer_call_and_return_conditional_losses_104246`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3a_activation_layer_call_and_return_conditional_losses_139358`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_wrapped_model_83162`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_wrapped_model_83162`) with ops

11 likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3c_se_reduce_layer_call_and_return_conditional_losses_100477`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block4b_expand_activation_layer_call_and_return_conditional_losses_102046`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3e_expand_activation_layer_call_and_return_conditional_losses_141384`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block4f_activation_layer_call_and_return_conditional_losses_103898`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6g_se_reduce_layer_call_and_return_conditional_losses_110440`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block7c_se_reduce_layer_call_and_return_conditional_losses_155753`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6h_activation_layer_call_and_return_conditional_losses_153548`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block7b_activation_layer_call_and_return_conditional_losses_155160`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block2c_expand_activation_layer_call_and_return_conditional_losses_98190`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3a_expand_activation_layer_call_and_return_conditional_losses_139219`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6h_activation_layer_call_and_return_conditional_losses_110833`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3e_activation_layer_call_and_return_conditional_losses_101310`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block5e_expand_activation_layer_call_and_return_conditional_losses_106393`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block6d_activation_layer_call_and_return_conditional_losses_151336`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_top_activation_layer_call_and_return_conditional_losses_112888`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block2d_activation_layer_call_and_return_conditional_losses_138252`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block1a_se_reduce_layer_call_and_return_conditional_losses_135374`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block5f_expand_activation_layer_call_and_return_conditional_losses_106833`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block3a_se_reduce_layer_call_and_return_conditional_losses_139398`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block2b_se_reduce_layer_call_and_return_conditional_losses_97889`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block1b_activation_layer_call_and_return_conditional_losses_96682`) with ops with custom gradients. Will likely fail if a gradient is requested.

WARNING:absl:Importing a function (`__inference_block5d_expand_activation_layer_call_and_return_conditional_losses_147373`) with ops with custom gradients. Will


```
1 likely fail if a gradient is requested.
WARNING:absl:Importing a function (_inference_block6a_expand_activation_layer
_call_and_return_conditional_losses_149585) with ops with custom gradients. Wi
ll likely fail if a gradient is requested.
WARNING:absl:Importing a function (_inference_block5c_se_reduce_layer_call_an
d_return_conditional_losses_146999) with ops with custom gradients. Will likel
y fail if a gradient is requested.
WARNING:absl:Importing a function (_inference_block2b_expand_activation_layer
_call_and_return_conditional_losses_97750) with ops with custom gradients. Wil
l likely fail if a gradient is requested.
WARNING:absl:Importing a function (_inference_block7c_se_reduce_layer_call_an
d_return_conditional_losses_112587) with ops with custom gradients. Will likel
y fail if a gradient is requested.
WARNING:absl:Importing a function (_inference_block7a_activation_layer_call_a
nd_return_conditional_losses_111713) with ops with custom gradients. Will like
ly fail if a gradient is requested.
WARNING:absl:Importing a function (_inference_block5a_expand_activation_layer
_call_and_return_conditional_losses_145761) with ops with custom gradients. Wi
ll likely fail if a gradient is requested.
```

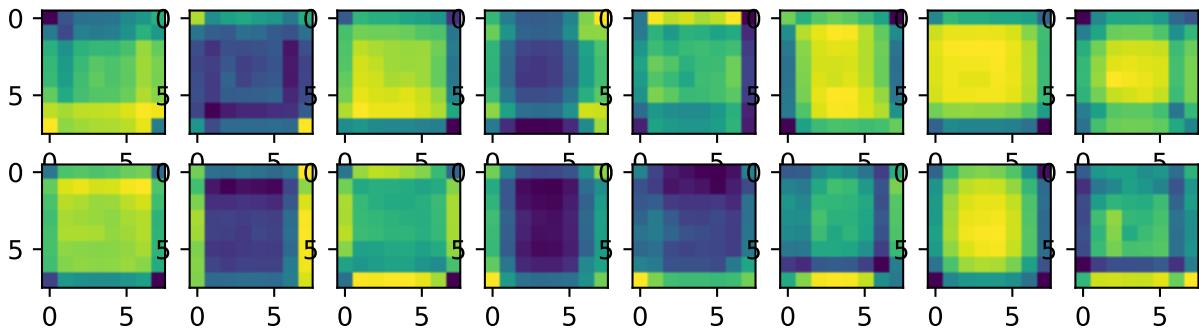
```
In [88]:  
preds = reload_model_1.predict(xi)  
print('predicted', preds)
```

```
predicted [[31.928402]]
```

```
In [96]:  
#feature maps of EfficientNet  
layer_names = ['block4d_expand_conv']  
layer_outputs = [reload_model_2.get_layer(layer_name).output for layer_name in  
activation_model = keras.models.Model(inputs=reload_model_2.input,outputs=layer  
activations = activation_model.predict(xi)  
plt.figure(figsize=(8,2))  
for i in range(16):  
    plt.subplot(2,8,i+1)  
    plt.imshow(activations[0,:,:,:,i],cmap='viridis')  
plt.show()
```

```
WARNING:tensorflow:10 out of the last 11 calls to <function Model.make_predict
_function.<locals>.predict_function at 0x0000016E1130D318> triggered tf.functi
on retracing. Tracing is expensive and the excessive number of tracings could
be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors
with different shapes, (3) passing Python objects instead of tensors. For (1),
please define your @tf.function outside of the loop. For (2), @tf.function has
experimental_relax_shapes=True option that relaxes argument shapes that can av
oid unnecessary retracing. For (3), please refer to https://www.tensorflow.or
g/guide/function#controlling_retracing and https://www.tensorflow.org/api_doc
s/python/tf/function for more details.
```

```
WARNING:tensorflow:10 out of the last 11 calls to <function Model.make_predict
_function.<locals>.predict_function at 0x0000016E1130D318> triggered tf.functi
on retracing. Tracing is expensive and the excessive number of tracings could
be due to (1) creating @tf.function repeatedly in a loop, (2) passing tensors
with different shapes, (3) passing Python objects instead of tensors. For (1),
please define your @tf.function outside of the loop. For (2), @tf.function has
experimental_relax_shapes=True option that relaxes argument shapes that can av
oid unnecessary retracing. For (3), please refer to https://www.tensorflow.or
g/guide/function#controlling_retracing and https://www.tensorflow.org/api_doc
s/python/tf/function for more details.
```



4. Try some new methods(Image and metadata fusion training)

4.1 Data reprocessing

```
In [ ]: csv_path = "../input/petfinder-pawpularity-score/"
model_path = "../input/pawpularitymodelfiles/"
image_path = "../input/petfinder-pawpularity-score/train/"
```

Calculate the VIF

VIF is a measure of the severity of complex (multiple) collinearity in multiple linear regression models. It represents the ratio of the variance of the estimator of the regression coefficient to that of the assumed non-linear correlation between the variances.

```
In [9]: from statsmodels.stats.outliers_influence import variance_inflation_factor
dftrainshow = dftrainshow.drop(['Id'], axis=1)
vif=[variance_inflation_factor(dftrainshow.values, dftrainshow.columns.get_loc
vif
```

```
Out[9]: [1.048306014529396,
10.118932122172204,
14.536468794682376,
6.260320912838975,
1.0105320415396544,
1.0917751935727698,
1.1934978760454777,
1.45446582391609,
2.0656144966285357,
2.0743070383688305,
1.4128104034445799,
1.5956617323271398,
3.7314179480481915]
```

We can find that some metadata features have very high VIF (measure of multicollinearity)

We can eliminate these features to achieve the purpose of data dimension reduction

```
In [10]: # Documentation
# Metadata features with VIF < 6
dense_features = ['Subject Focus', 'Action', 'Accessory', 'Group', 'Collage', '
```

```
In [15]: # Documentation
# some initialization
```

```

numsplit = 5
fold = [0,1,2,3,4]
batch_size = 64

# Documentation
# Considering this is a regression problem, we use the sigmoid function
def sigmoid(x):
    return 1 / (1 + math.exp(-x))

```

Process Image duplication(Some pictures are duplicate and need to be removed.)

In [26]:

```

from imagededup.methods import PHash
from imagededup.utils import plot_duplicates

phasher = PHash()
encodings = phasher.encode_images(image_dir=r'../input/petfinder-pawpularity-DuplicateImgaes')
DuplicateImgaes = phasher.find_duplicates(encoding_map=encodings)

```

```

2021-12-02 13:19:56,010: INFO Start: Calculating hashes...
100%|██████████| 9912/9912 [02:23<00:00, 69.20it/s]
2021-12-02 13:22:19,464: INFO End: Calculating hashes!
2021-12-02 13:22:19,468: INFO Start: Evaluating hamming distances for getting duplicates
2021-12-02 13:22:19,469: INFO Start: Retrieving duplicates using Cython Brute force algorithm
100%|██████████| 9912/9912 [00:00<00:00, 15935.55it/s]
2021-12-02 13:22:20,336: INFO End: Retrieving duplicates using Cython Brute force algorithm
2021-12-02 13:22:20,337: INFO End: Evaluating hamming distances for getting duplicates

```

Show repetitive images

In [27]:

```

plot_duplicates(image_dir=r'../input/petfinder-pawpularity-score/train/',
                 duplicate_map=DuplicateImgaes,
                 filename='3877f2981e502fe1812af38d4f511fd2.jpg')

```

Original Image: 3877f2981e502fe1812af38d4f511fd2.jpg



Generate a data set of duplicate images

In [28]:

```

repeat_img = []
is_img = []

for k, v in DuplicateImgaes.items():
    if not v:
        is_img.append(k)
    elif k not in repeat_img:

```

```

        is_img.append(k)
        repeat_img.extend(v)
    else:
        repeat_img.extend(v)

duplicates = []
for k in repeat_img:
    m = k.split('.')
    duplicates.append(m[0])
print(duplicates)
print(len(duplicates))

```

```

['e09a818b7534422fb4c688f12566e38f', '685f7da30783c5c77406bde3987a5db1', '9a0238499efb15551f06ad583a6fa951', '0d6abba9e81300164592ab9a66c52f16', 'dbf25ce0b2a5d3cb43af95b2bd855718', 'd6bf1c194e3d459f2569f250fc1d5db4', 'c504568822c53675a4f425c8e5800a36', '5a642ecc14e9c57a05b8e010414011f2', 'bf8501acaeee dc2a421bac3d9af58bb7', '01430d6ae02e79774b651175edd40842', '776d92567e4d173093527335741f49c6', '1feb99c2a4cac3f3c4f8a4510421d6f5', '6ae42b731c00756ddd291fa615c822a1', '13d215b4c71c3dc603cd13fc3ec80181', '902786862cbae94e890a090e5700298b', '67e97de8ec7ddcda59a58b027263cdcc', '2ddbb3990b65011a3d46c58edcf6dca3', '988b31dd48a1bc867dbc9e14d21b05f6', '68e55574e523cf1cdc17b60ce6cc2f60', '82e51d0ead705357f211e589a1efecbe', 'e359704524fa26d6a3dc8bfeeae2e', '72b33c9c368d86648b756143ab19baeb', '2b737750362ef6b31068c4a4194909ed', '64ef283f60d9a1a10545444731d5f6ce', '1673ded6827ae62f2507734e04ed52ed', '871bb3cbdf48bd3bfd5a6779e752613e', 'b5450c0ee3436529bcefc60efc7bc26a', '5ef7ba98fc97917aec56ded5d5c2b099', '3877f2981e502fe1812af38d4f511fd2', '2bb0677b6e5024f8e374bf9e288f137b', '43bd09ca68b3bc2b0c549fd309d1ba', '151c5ea5345c25e4d1f92a441b586b1a', 'ecd0afc94bc55972ce058ccb16e3045e', '851c7427071af2eaf38af0def360987', '306e4816eb2e3427008ea48389ebf859', '9f5a457ce7e22eecd0992f4ea17b6107', 'ccc822df80e2c9adb4a68e56d68d97c5', '24b10e866f289f8e7f7686e1e77a1235', '6322aa006010abfa543a3ac224bb3d01', 'b148cbea87c3dcc65a05b15f78910715', '08440f8c2c040cf2941687de6dc5462f', '43ab682adde9c14adb7c05435e5f2e0e', '2757c5b3109466cb11bad0e5bd2919d2', 'dd042410dc7f02e648162d7764b50900', 'ecebf6a4483edc92e210da9e3e20f1ec', 'a9513f7f0c93e179b87c01be847b3e4c', '6dc1ae625a3bfb50571efedc0afc297c', 'cc66d7a6a006c15e3ee3307927cc5a02', 'bca6811ee0a78bdcc41b659624608125', 'd050e78384bd8b20e7291b3efedf6a5b', '9b3267c1652691240d78b7b3d072baf3', '0e94c5206de8777a57b63a4a52a751ef', 'fea5968d4979d8d98d964c31e0c7fe66', '5da97b511389a1b62ef7a55b0a19a532', '8ffd3ae7ab3726cff7ca28697687a42', '99db12809c7672aa9b1c7954cd65ef1f', '9466b0b218c168a2417231c25a75dbae', '373c763f5218610e9b3f82b12ada8ae5', '9d41b2d3779ac2ae10d3398d14cd5c9a', '41c85c2c974cc15ca77f5ababb652f84', '03d82e64d1b4d99f457259f03ebe604d', '83f213c33ad2fbf2a59e57529ee6f0e0', '264845a4236bc9b95123dde3fb809a88', '14792e17a677099449059d58a507a42c', '38426ba3cbf5484555f2b5e9504a6b03', 'b49ad3aac4296376d7520445a27726de', 'ba39a25e1c28ef319e4a2cbbe41d2dd0', '1059231cf2948216fcc2ac6afb4f8db8', '763d66b9cf01069602a968e573feb334', '54563ff51aa70ea8c6a9325c15f55399', 'cc1306f5b447ea83496e10d8f77509a2', '0c4d454d8f09c90c655bd0e2af6eb2e5', '5a5c229e1340c0da7798b26edf86d180', 'dbc47155644aeb3edd1bd39dba9b6953', '555e9ac5c3620688ce989072334420ee', '376fec474c48882e5790b21c6c55de98', 'c25384f6d93ca6b802925da84dfa453e', 'b967656eb7e648a524ca4ffbbc172c06', '16b16c71d71bd20843683feebc3af356', 'b956edfd0677dd6d95de6cb29a85db9c', '92740e4e44c2572a43199931789e7166', '87c6a8f85af93b84594a36f8ffd5d6b8', '303e1a5607e28c4ab68a6a39be3e9bae', 'fe47539e989df047507eaa60a16bc3fd', 'b86589c3e85f784a5278e377b726a4d4', '78a02b3cb6ed38b2772215c0c0a7f78e', '7d6c4c2c3889b7c24843042a4d907c7c', '589286d5bfdc1b26ad0bf7d4b7f74816', '6cb18e0936faa730077732a25c3dfb94', 'cd909abf8f425d7e646eebe4d3bf4769']
90

```

Remove duplicate images

In [13]:

```

def rid_duplicates(df):
    for id in duplicates:
        df = df.drop(df[df.Id == id].index)
    return df.reset_index(drop=True)

# Documentation
# Data set sharding is used for Kfold cross-validation
def create_folds(data, num_splits):
    data["kfold"] = -1

```

```

num_bins = int(np.floor(1 + np.log2(len(data))))
data.loc[:, "bins"] = pd.cut(data["Pawpularity"], bins=num_bins, labels=False)
kf = model_selection.StratifiedKFold(n_splits=num_splits, shuffle=True, random_state=42)
for f, (t_, v_) in enumerate(kf.split(X=data, y=data.bins.values)):
    data.loc[v_, 'kfold'] = f
data = data.drop("bins", axis=1)
return data

```

The data set was divided into 5 points for cross-validation

```
In [19]: df_test = pd.read_csv(csv_path+"test.csv")
df = create_folds(rid_duplicates(pd.read_csv(csv_path+"train.csv")), num_splits)
dftrainshow = df[df.kfold == 1].reset_index(drop=True)
dftrainshow.head()
```

Out[19]:

		Id	Subject Focus	Eyes	Face	Near	Action	Accessory	Group
0	001dc955e10590d3ca4673f034feef2		0	0	0	1	0	0	1
1	005017716086b8d5e118dd9fe26459b1		0	1	1	1	0	0	0
2	00768659c1c90409f81dcdecbd270513		0	1	1	0	0	0	0
3	0095f81bab3b68a4f70e99f0fcec7b06		0	0	1	1	0	0	0
4	00eb7cc3d158a6addbcbde1f267d7d15		0	1	1	1	0	0	0

In [14]:

```

# Documentation
# Data enhancement
# Albumentation is a lib for image augmentation operations.

train_aug = albumentations.Compose(
    [
        albumentations.Resize(128, 128, p=1),
        albumentations.HueSaturationValue(hue_shift_limit=0.2, sat_shift_limit=0.2, val_shift_limit=0.2, p=1),
        albumentations.RandomBrightnessContrast(brightness_limit=(-0.1, 0.1), contrast_limit=(-0.1, 0.1), p=1),
        albumentations.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.224], p=1),
    ],
    p=1.0,
)

valid_aug = albumentations.Compose(
    [
        albumentations.Resize(128, 128, p=1),
        albumentations.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.224], p=1),
    ],
    p=1.0,
)

```

4.2 New feature representation and custom neural network

Since the data is multimodal, we need to perform data fusion.

We use PyTorch to build and customize networks

We use an external EfficientNetB0 model in the hope of better performance

We consider to train the image to extract features first, then link the image features and metadata features in the full connection layer to obtain new feature representation, and

finally train the new joint features to produce output

We use GPU to accelerate the model training process

In [16]:

```
# @Documentation
# use PyTorch to build and customize networks
# Class to represent dataset's image , metadata and targets , rolled into 1 object
# Returns dict of torch tensors form of image, metadata & targets
class PawpularDataset:
    def __init__(self, image_paths, dense_features, targets, augmentations):
        self.image_paths = image_paths
        self.dense_features = dense_features
        self.targets = targets
        self.augmentations = augmentations

    def __len__(self):
        return len(self.image_paths)

    def __getitem__(self, item):
        image = cv2.imread(self.image_paths[item])
        image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
        if self.augmentations is not None:
            augmented = self.augmentations(image=image)
            image = augmented["image"]
        image = np.transpose(image, (2, 0, 1)).astype(np.float32)
        features = self.dense_features[item, :]
        targets = self.targets[item]
        return {
            "image": torch.tensor(image, dtype=torch.float),
            "features": torch.tensor(features, dtype=torch.float),
            "targets": torch.tensor(targets, dtype=torch.float),
        }

# @Documentation
# Model definition
# firstly send an image into a model to extract image features
# dense1 concatenate image features & metadata features
# The features of the union are 128+9 dimensions, where 9 is the metadata features
class PawpularModel(tez.Model):
    def __init__(self, get_pretrained, model_string):
        super().__init__()
        self.get_pretrained = get_pretrained
        self.model_string = model_string
        self.model = timm.create_model(self.model_string, pretrained=self.get_pretrained)
        self.model.classifier = nn.Linear(self.model.classifier.in_features, 128)
        self.dropout = nn.Dropout(0.1)
        self.dense1 = nn.Linear(128+len(dense_features), 64)
        self.dense2 = nn.Linear(64, 1)
        self.step_scheduler_after = "epoch"

    def monitor_metrics(self, outputs, targets):
        outputs = outputs.cpu().detach().numpy()
        targets = targets.cpu().detach().numpy()
        rmse = metrics.mean_squared_error(targets, outputs, squared=False)
        return {"rmse": rmse}

    def fetch_scheduler(self):
        sch = torch.optim.lr_scheduler.CosineAnnealingWarmRestarts(
            self.optimizer, T_0=10, T_mult=1, eta_min=1e-6, last_epoch=-1
        )
        return sch
```

```

def fetch_optimizer(self):
    opt = torch.optim.Adam(self.parameters(), lr=1e-4)
    return opt

def forward(self, image, features, targets=None):
    x = self.model(image)
    x = self.dropout(x)
    x = torch.cat([x, features], dim=1)
    x = self.dense1(x)
    x = self.dense2(x)
    if targets is not None:
        loss = nn.MSELoss()(x, targets.view(-1, 1))
        metrics = self.monitor_metrics(x, targets)
        return x, loss, metrics

    return x, 0, {}

```

In [17]:

```

import tensorflow as tf
from tensorflow import keras
from tensorflow.keras.models import Sequential, Model
from tensorflow.keras.layers import Dense, Activation, Conv2D, Flatten, Dropout
from tensorflow.keras import backend as K
nn_model_name = 'tf_efficientnet_b0_ns'
nn_model_filename = "nn_" + nn_model_name

```

4.3 Model training(K-fold Cross validation)

K-fold Cross validation

In [20]:

```

model = PawpularModel(True, nn_model_name)
for i in fold:
    nn_model_Fname = nn_model_filename+f"_f{i}.bin"
    df_train = df[df.kfold != i].reset_index(drop=True)
    df_valid = df[df.kfold == i].reset_index(drop=True)
    train_img_paths = [image_path+f"{x}.jpg" for x in df_train["Id"].values]
    valid_img_paths = [image_path+f"{x}.jpg" for x in df_valid["Id"].values]
    train_dataset = PawpularDataset(
        image_paths=train_img_paths,
        dense_features=df_train[dense_features].values,
        targets=df_train.Pawpularity.values/100.0,
        augmentations=train_aug,
    )
    valid_dataset = PawpularDataset(
        image_paths=valid_img_paths,
        dense_features=df_valid[dense_features].values,
        targets=df_valid.Pawpularity.values/100.0,
        augmentations=valid_aug,
    )

    es = EarlyStopping(monitor="valid_rmse", model_path= nn_model_Fname,
                       patience=3, mode="min", save_weights_only=True,)

    model.fit(train_dataset, valid_dataset=valid_dataset, train_bs=batch_size,
              valid_bs=2*batch_size, epochs=30, callbacks=[es], fp16=True, )

```

/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

```
    cpuset_checked))
100%|██████████| 123/123 [01:29<00:00,  1.38it/s, loss=0.0512, rmse=0.2
23, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:20<00:00,  1.26s/it, loss=0.0422, rmse=0.20
5, stage=valid]
Validation score improved (inf --> 0.20490560308098793). Saving model!
 0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.49it/s, loss=0.0323, rmse=0.1
79, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.22s/it, loss=0.0421, rmse=0.20
5, stage=valid]
EarlyStopping counter: 1 out of 3
 0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.50it/s, loss=0.0218, rmse=0.1
47, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:20<00:00,  1.27s/it, loss=0.0449, rmse=0.21
1, stage=valid]
EarlyStopping counter: 2 out of 3
 0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.0141, rmse=0.1
18, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
```

```
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.23s/it, loss=0.046, rmse=0.214,
stage=valid]
EarlyStopping counter: 3 out of 3
    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.0102, rmse=0.
1, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.24s/it, loss=0.0479, rmse=0.21
8, stage=valid]
Validation score improved (inf --> 0.21823222376406193). Saving model!
    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.00757, rmse=0.
0864, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.24s/it, loss=0.0472, rmse=0.21
7, stage=valid]
Validation score improved (0.21823222376406193 --> 0.21669852267950773). Savin
g model!
    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.50it/s, loss=0.00653, rmse=0.
0803, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
```

be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

cpuset_checked))

100%|██████████| 16/16 [00:19<00:00, 1.24s/it, loss=0.0472, rmse=0.21
7, stage=valid]

EarlyStopping counter: 1 out of 3

0%| | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

cpuset_checked))

100%|██████████| 123/123 [01:23<00:00, 1.48it/s, loss=0.00592, rmse=0.0762, stage=train]

0%| | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

cpuset_checked))

100%|██████████| 16/16 [00:19<00:00, 1.23s/it, loss=0.0467, rmse=0.216, stage=valid]

Validation score improved (0.21669852267950773 --> 0.21550152450799942). Saving model!

0%| | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

cpuset_checked))

100%|██████████| 123/123 [01:24<00:00, 1.46it/s, loss=0.00531, rmse=0.0722, stage=train]

0%| | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

cpuset_checked))

100%|██████████| 16/16 [00:20<00:00, 1.27s/it, loss=0.0469, rmse=0.216, stage=valid]

EarlyStopping counter: 1 out of 3

0%| | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

cpuset_checked))

100%|██████████| 123/123 [01:28<00:00, 1.39it/s, loss=0.00512, rmse=0.0708, stage=train]

0%| | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

```
    cpuset_checked))
100%|██████████| 16/16 [00:20<00:00,  1.30s/it, loss=0.0468, rmse=0.21
6, stage=valid]
EarlyStopping counter: 2 out of 3

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:29<00:00,  1.37it/s, loss=0.00583, rmse=0.
0758, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:21<00:00,  1.36s/it, loss=0.0492, rmse=0.22
1, stage=valid]
EarlyStopping counter: 3 out of 3

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:29<00:00,  1.37it/s, loss=0.00659, rmse=0.
0805, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:21<00:00,  1.32s/it, loss=0.0463, rmse=0.21
4, stage=valid]
Validation score improved (inf --> 0.21446009073406458). Saving model!

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:28<00:00,  1.39it/s, loss=0.00597, rmse=0.
0764, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:20<00:00,  1.31s/it, loss=0.0471, rmse=0.21
6, stage=valid]
EarlyStopping counter: 1 out of 3
```

```
0% | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:25<00:00,  1.43it/s, loss=0.0052, rmse=0.0
715, stage=train]
0% | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:20<00:00,  1.31s/it, loss=0.0447, rmse=0.21
1, stage=valid]
Validation score improved (0.21446009073406458 --> 0.21071808971464634). Saving model!
0% | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:28<00:00,  1.39it/s, loss=0.00429, rmse=0.
0648, stage=train]
0% | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:21<00:00,  1.33s/it, loss=0.0442, rmse=0.20
9, stage=valid]
Validation score improved (0.21071808971464634 --> 0.20949057023972273). Saving model!
0% | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:26<00:00,  1.42it/s, loss=0.00386, rmse=0.
0614, stage=train]
0% | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.23s/it, loss=0.0444, rmse=0.21,
stage=valid]
EarlyStopping counter: 1 out of 3
0% | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
```

worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

```
    cpuset_checked))
100%|██████████| 123/123 [01:25<00:00,  1.45it/s, loss=0.00351, rmse=0.
0584, stage=train]
  0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.21s/it, loss=0.044, rmse=0.209,
stage=valid]
EarlyStopping counter: 2 out of 3
  0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.49it/s, loss=0.00325, rmse=0.
0562, stage=train]
  0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:20<00:00,  1.25s/it, loss=0.0441, rmse=0.20
9, stage=valid]
EarlyStopping counter: 3 out of 3
  0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.49it/s, loss=0.00301, rmse=0.
0542, stage=train]
  0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.24s/it, loss=0.0437, rmse=0.20
8, stage=valid]
Validation score improved (inf --> 0.2082573790103197). Saving model!
  0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary
```

```
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.00288, rmse=0.
0529, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.25s/it, loss=0.0436, rmse=0.20
8, stage=valid]
EarlyStopping counter: 1 out of 3

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.00344, rmse=0.
0579, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.25s/it, loss=0.044, rmse=0.209,
stage=valid]
EarlyStopping counter: 2 out of 3

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.00432, rmse=0.
0648, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.23s/it, loss=0.0431, rmse=0.20
7, stage=valid]
Validation score improved (0.2082573790103197 --> 0.20686230156570673). Saving
model!

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.49it/s, loss=0.00401, rmse=0.
```

```
0625, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:18<00:00,  1.16s/it, loss=0.0433, rmse=0.20
7, stage=valid]
EarlyStopping counter: 1 out of 3

 0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.49it/s, loss=0.00357, rmse=0.
0588, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.20s/it, loss=0.0431, rmse=0.20
7, stage=valid]
EarlyStopping counter: 2 out of 3

 0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.50it/s, loss=0.00344, rmse=0.
0577, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.20s/it, loss=0.0423, rmse=0.20
5, stage=valid]
Validation score improved (0.20686230156570673 --> 0.20489057339727879). Saving model!

 0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.00275, rmse=0.
0519, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
```

worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

```
    cpuset_checked))
100%|██████████| 16/16 [00:18<00:00,  1.18s/it, loss=0.0425, rmse=0.20
6, stage=valid]
EarlyStopping counter: 1 out of 3
    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.50it/s, loss=0.00266, rmse=0.
0507, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.20s/it, loss=0.0424, rmse=0.20
5, stage=valid]
EarlyStopping counter: 2 out of 3
    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.0023, rmse=0.0
474, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:18<00:00,  1.17s/it, loss=0.0417, rmse=0.20
3, stage=valid]
Validation score improved (0.20489057339727879 --> 0.20348703395575285). Saving model!
    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.51it/s, loss=0.00222, rmse=0.
0464, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
```

```
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.25s/it, loss=0.0419, rmse=0.20
4, stage=valid]
EarlyStopping counter: 1 out of 3

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:20<00:00,  1.52it/s, loss=0.00214, rmse=0.
0457, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:20<00:00,  1.26s/it, loss=0.0419, rmse=0.20
4, stage=valid]
EarlyStopping counter: 2 out of 3

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:20<00:00,  1.52it/s, loss=0.00321, rmse=0.
0545, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.24s/it, loss=0.0424, rmse=0.20
5, stage=valid]
EarlyStopping counter: 3 out of 3

    0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:21<00:00,  1.52it/s, loss=0.00365, rmse=0.
0595, stage=train]
    0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
```

```
100%|██████████| 16/16 [00:20<00:00, 1.25s/it, loss=0.0424, rmse=0.20
5, stage=valid]
Validation score improved (inf --> 0.20516619179397821). Saving model!
0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
cpuset_checked))
100%|██████████| 123/123 [01:22<00:00, 1.50it/s, loss=0.00347, rmse=0.
0578, stage=train]
0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
cpuset_checked))
100%|██████████| 16/16 [00:19<00:00, 1.24s/it, loss=0.0425, rmse=0.20
5, stage=valid]
EarlyStopping counter: 1 out of 3
0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
cpuset_checked))
100%|██████████| 123/123 [01:22<00:00, 1.49it/s, loss=0.00303, rmse=0.
0542, stage=train]
0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
cpuset_checked))
100%|██████████| 16/16 [00:19<00:00, 1.20s/it, loss=0.042, rmse=0.204,
stage=valid]
EarlyStopping counter: 2 out of 3
0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/
torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
cpuset_checked))
100%|██████████| 123/123 [01:22<00:00, 1.48it/s, loss=0.00279, rmse=0.
0519, stage=train]
0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
cpuset_checked))
100%|██████████| 16/16 [00:19<00:00, 1.21s/it, loss=0.0418, rmse=0.20
4, stage=valid]
Validation score improved (0.20516619179397821 --> 0.20379382465034723). Savin
g model!
```

```
0% | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.49it/s, loss=0.00257, rmse=0.0498, stage=train]
0% | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.21s/it, loss=0.0416, rmse=0.203, stage=valid]
EarlyStopping counter: 1 out of 3
0% | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.48it/s, loss=0.00237, rmse=0.0478, stage=train]
0% | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.19s/it, loss=0.0415, rmse=0.203, stage=valid]
EarlyStopping counter: 2 out of 3
0% | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 123/123 [01:23<00:00,  1.47it/s, loss=0.00202, rmse=0.044, stage=train]
0% | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.22s/it, loss=0.0412, rmse=0.202, stage=valid]
Validation score improved (0.20379382465034723 --> 0.2023298181593418). Saving model!
0% | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
```

m is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

```

    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.50it/s, loss=0.00181, rmse=0.
0419, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:20<00:00,  1.25s/it, loss=0.0413, rmse=0.20
3, stage=valid]
EarlyStopping counter: 1 out of 3

 0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.50it/s, loss=0.00175, rmse=0.
0411, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.24s/it, loss=0.0414, rmse=0.20
3, stage=valid]
EarlyStopping counter: 2 out of 3

 0%|          | 0/123 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 123/123 [01:22<00:00,  1.50it/s, loss=0.00226, rmse=0.
0464, stage=train]
 0%|          | 0/16 [00:00<?, ?it/s]/opt/conda/lib/python3.7/site-packages/t
orch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create 8
worker processes in total. Our suggested max number of worker in current syste
m is 2, which is smaller than what this DataLoader is going to create. Please
be aware that excessive worker creation might get DataLoader running slow or e
ven freeze, lower the worker number to avoid potential slowness/freeze if nece
ssary.
    cpuset_checked))
100%|██████████| 16/16 [00:19<00:00,  1.25s/it, loss=0.0411, rmse=0.20
2, stage=valid]
EarlyStopping counter: 3 out of 3

```

In [21]:

```

allfolds_oof_nnreg_preds = []
allfolds_oof_validation_tgts = []
allfolds_submit_nnreg_preds = []
test_img_paths = [csv_path+"test/"+f"{x}.jpg" for x in df_test["Id"].values]

```

4.4 Model prediction and error analysis

We use OOF(Out of fold predictions) to calculate the RMSE

We have previously generated five folding models through cross-training, and now we want to predict and calculate RMSE for the validation set data by using these five models separately

```
In [22]: nn_model_filename = "nn_" + nn_model_name
for i in fold:
    model = PawpularModel(False, nn_model_name)
    nn_model_Fname = nn_model_filename + f"_f{i}.bin"
    df_train = df[df.kfold != i].reset_index(drop=True)
    df_valid = df[df.kfold == i].reset_index(drop=True)
    train_img_paths = [image_path + f"{x}.jpg" for x in df_train["Id"].values]
    valid_img_paths = [image_path + f"{x}.jpg" for x in df_valid["Id"].values]

    model.load(nn_model_Fname, device="cuda", weights_only=True)

    valid_dataset = PawpularDataset(
        image_paths=valid_img_paths,
        dense_features=df_valid[dense_features].values,
        targets=df_valid['Pawpularity'].values / 100.00,
        augmentations=valid_aug)
    valid_predictions = model.predict(valid_dataset, batch_size=2 * batch_size,
                                      n_jobs=4)

    sglfold_oof_nnreg_pred = []
    for preds in valid_predictions:
        sglfold_oof_nnreg_pred.extend(preds[:, :1].ravel().tolist())

    sglfold_oof_nnreg_pred = [sigmoid(x) * 100 for x in sglfold_oof_nnreg_pred]
    allfolds_oof_nnreg_preds.append(sglfold_oof_nnreg_pred)

    sglfold_oof_validation_tgts = df_valid['Pawpularity'].values
    allfolds_oof_validation_tgts.append(sglfold_oof_validation_tgts)

    rsme = np.sqrt(np.mean((allfolds_oof_validation_tgts[-1] - np.array(allfolds_oof_nnreg_preds)) ** 2))
    print(f"\nNN RSME = {rsme}\n")

    test_dataset = PawpularDataset(
        image_paths=test_img_paths,
        dense_features=df_test[dense_features].values,
        targets=np.ones(len(test_img_paths)),
        augmentations=valid_aug)
    test_predictions = model.predict(test_dataset, batch_size=2 * batch_size, n_jobs=4)

    sglfold_submit_nnreg_pred = []
    embed = np.array([]).reshape((0, 128 + len(dense_features)))
    for preds in test_predictions:
        sglfold_submit_nnreg_pred.extend(preds[:, :1].ravel().tolist())

    sglfold_submit_nnreg_pred = [sigmoid(x) * 100 for x in sglfold_submit_nnreg_pred]
    allfolds_submit_nnreg_preds.append(sglfold_submit_nnreg_pred)
    print(allfolds_submit_nnreg_preds)
```

```
100%|██████████| 16/16 [00:18<00:00,  1.16s/it, stage=test]
NN RSME = 29.62849831350185
```

```
100%|██████████| 1/1 [00:00<00:00,  4.49it/s, stage=test]
[[69.20652005974578, 66.60076046726255, 68.32211126996289, 67.55263450672979,
 70.73887364407622, 67.00942373994845, 70.43337820781808, 68.05879750655578]]
```

```
100%|██████████| 16/16 [00:19<00:00,  1.20s/it, stage=test]
NN RSME = 26.180267674950866
```

```
100% |██████████| 1/1 [00:00<00:00, 4.76it/s, stage=test]
[[69.20652005974578, 66.60076046726255, 68.32211126996289, 67.55263450672979,
70.73887364407622, 67.00942373994845, 70.43337820781808, 68.05879750655578],
[73.92659047025671, 72.70252822823096, 72.73608870895325, 73.79345390234249, 7
4.09885931434309, 72.50133387684623, 74.39376139719474, 72.82914980848496]]
100% |██████████| 16/16 [00:18<00:00, 1.17s/it, stage=test]
NN RSME = 25.90147198154856

100% |██████████| 1/1 [00:00<00:00, 2.95it/s, stage=test]
[[69.20652005974578, 66.60076046726255, 68.32211126996289, 67.55263450672979,
70.73887364407622, 67.00942373994845, 70.43337820781808, 68.05879750655578],
[73.92659047025671, 72.70252822823096, 72.73608870895325, 73.79345390234249, 7
4.09885931434309, 72.50133387684623, 74.39376139719474, 72.82914980848496], [7
1.17595320935692, 69.9634624897766, 70.34557844536113, 70.92451886458613, 71.6
143214021789, 69.99064495557552, 71.25155097507741, 70.27943117724344], [70.32
467906732232, 69.51264806246613, 69.63291523615496, 70.26848905322967, 70.3399
0331245224, 69.40752749669211, 70.44348721090891, 69.5242857127037]]
100% |██████████| 16/16 [00:18<00:00, 1.18s/it, stage=test]
NN RSME = 26.153531144751692

100% |██████████| 1/1 [00:00<00:00, 4.92it/s, stage=test]
[[69.20652005974578, 66.60076046726255, 68.32211126996289, 67.55263450672979,
70.73887364407622, 67.00942373994845, 70.43337820781808, 68.05879750655578],
[73.92659047025671, 72.70252822823096, 72.73608870895325, 73.79345390234249, 7
4.09885931434309, 72.50133387684623, 74.39376139719474, 72.82914980848496], [7
1.17595320935692, 69.9634624897766, 70.34557844536113, 70.92451886458613, 71.6
143214021789, 69.99064495557552, 71.25155097507741, 70.27943117724344], [70.32
467906732232, 69.51264806246613, 69.63291523615496, 70.26848905322967, 70.3399
0331245224, 69.40752749669211, 70.44348721090891, 69.5242857127037], [69.54918
699851214, 68.80791839612353, 69.12994806032955, 69.47949981318156, 69.4388702
984862, 68.90416488946236, 69.66024444368028, 68.87413211723481]]
```

We use "128+9" joint features for feature representation (where 128 is the image features and 9 is the metadata features). I think we can do dimension reduction or weight assignment on joint features, because the image occupies a relatively high dimension. Metadata features are also very important considerations and should be given more attention

I think we still have more space for optimization in the model. In the processing of the data set, we deleted the larger eigenvalues of VIF. If time is sufficient, maybe we can consider taking all the eigenvalues into consideration for a training and comparative analysis. In some places the choice of parameters is not optimal.

```
In [ ]: true = np.hstack(allfolds_oof_validation_tgts)
oof = np.hstack(allfolds_oof_nnreg_preds)
rsme = np.sqrt( np.mean( (oof - true)**2.0 ) )
print('CNN RSME = ', rsme)
```

4.5 Submission

The prediction result is calculated according to the model of each fold, and the average value is taken at the end

In [24]:

```
allfolds_submit_nnreg_preds = np.mean(np.column_stack(allfolds_submit_nnreg_p:
df_test["Pawpularity"] = allfolds_submit_nnreg_preds
df_test = df_test[["Id", "Pawpularity"]]
df_test.to_csv("submission.csv", index=False)
```