OcularCNN-Improvements

In [1]:
```python
import pandas as pd
import random
import numpy as np

theFile = pd.read_csv("../data/full_df.csv")
theFile
```

Out[1]:

	ID	Patient Age	Patient Sex	Left-Fundus	Right-Fundus	Left-Diagnostic Keywords	Right-Diagnostic Keywords	N	D	G	C	A	H	M	O	filepath	labels
0	0	69	Female	0_left.jpg	0_right.jpg	cataract	normal fundus	0	0	0	1	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']
1	1	57	Male	1_left.jpg	1_right.jpg	normal fundus	normal fundus	1	0	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']
2	2	42	Male	2_left.jpg	2_right.jpg	laser spot, moderate non proliferative retinopathy	moderate non proliferative retinopathy	0	1	0	0	0	0	0	1	../input/ocular-disease-recognition-odir5k/ODI...	['D']
3	4	53	Male	4_left.jpg	4_right.jpg	macular epiretinal membrane	mild nonproliferative retinopathy	0	1	0	0	0	0	0	1	../input/ocular-disease-recognition-odir5k/ODI...	['D']
4	5	50	Female	5_left.jpg	5_right.jpg	moderate non proliferative retinopathy	moderate non proliferative retinopathy	0	1	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['D']
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
6387	4686	63	Male	4686_left.jpg	4686_right.jpg	severe nonproliferative retinopathy	proliferative diabetic retinopathy	0	1	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['D']
6388	4688	42	Male	4688_left.jpg	4688_right.jpg	moderate non proliferative retinopathy	moderate non proliferative retinopathy	0	1	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['D']
6389	4689	54	Male	4689_left.jpg	4689_right.jpg	mild nonproliferative retinopathy	normal fundus	0	1	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['D']
6390	4690	57	Male	4690_left.jpg	4690_right.jpg	mild nonproliferative retinopathy	mild nonproliferative retinopathy	0	1	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['D']
6391	4784	58	Male	4784_left.jpg	4784_right.jpg	hypertensive retinopathy, age-related macular d...	hypertensive retinopathy, age-related macular d...	0	0	0	0	1	1	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['H']

6392 rows × 19 columns

In [2]:
```python
theFile.drop(np.linspace(1, 3194, 3193, dtype = int), inplace = True )
target = theFile.loc[(theFile['Left-Diagnostic Keywords'] == 'cataract') | (theFile['Left-Diagnostic Ke
ywords'] == 'normal fundus')]



target = theFile.loc[(theFile['Left-Diagnostic Keywords'] == 'cataract') | (theFile['Left-Diagnostic Ke
ywords'] == 'normal fundus')]
```

In [3]:
```python
print(len(target.loc[target['Left-Diagnostic Keywords'] == 'cataract']))
print(len(target.loc[target['Left-Diagnostic Keywords'] == 'normal fundus']))
```

```
135
1433
```

In [4]:
```python
#As shown above, there are much more normal fundus entries than cataract
#Let's sample the normal fundus to get equivalent numbers of each class
cataract = target.loc[target['Left-Diagnostic Keywords'] == 'cataract']
normal = target.loc[target['Left-Diagnostic Keywords'] == 'normal fundus']
normal = normal.sample(135)
frames = [cataract, normal]
target = pd.concat(frames)

#now shuffle these randomly
# shuffle the DataFrame rows
target= target.sample(frac = 1)
```

In [5]:
```python
pictures = target[['filename']]
```

In [6]:
```python
target
```

Out[6]:

	ID	Patient Age	Patient Sex	Left-Fundus	Right-Fundus	Left-Diagnostic Keywords	Right-Diagnostic Keywords	N	D	G	C	A	H	M	O	filepath	labels	ta
4058	990	41	Female	990_left.jpg	990_right.jpg	normal fundus	macular epiretinal membrane	0	0	0	0	0	0	0	1	../input/ocular-disease-recognition-odir5k/ODI...	['N']	
6232	4496	53	Female	4496_left.jpg	4496_right.jpg	normal fundus	mild nonproliferative retinopathy	0	1	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']	
4132	1113	56	Female	1113_left.jpg	1113_right.jpg	normal fundus	mild nonproliferative retinopathy	0	1	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']	
4803	2432	45	Female	2432_left.jpg	2432_right.jpg	normal fundus	normal fundus	1	0	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']	
4588	2123	47	Male	2123_left.jpg	2123_right.jpg	cataract	refractive media opacity	0	0	0	1	0	0	0	1	../input/ocular-disease-recognition-odir5k/ODI...	['C']	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
5080	2740	55	Male	2740_left.jpg	2740_right.jpg	normal fundus	normal fundus	1	0	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']	
4574	2109	82	Male	2109_left.jpg	2109_right.jpg	cataract	cataract	0	0	0	1	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['C']	
4585	2120	72	Female	2120_left.jpg	2120_right.jpg	cataract	cataract	0	0	0	1	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['C']	
5243	2916	61	Male	2916_left.jpg	2916_right.jpg	normal fundus	normal fundus	1	0	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']	
4634	2172	73	Male	2172_left.jpg	2172_right.jpg	cataract	normal fundus	0	0	0	1	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['C']	

270 rows × 19 columns

In [7]:
```python
classnames = ['[\'N\']', '[\'D\']','[\'G\']','[\'C\']','[\'A\']','[\'H\']','[\'M\']','[\'O\']']
classnames
```

Out[7]:
```
["['N']", "['D']", "['G']", "['C']", "['A']", "['H']", "['M']", "['O']"]
```

In [8]:
```python
pictures = pictures.values.tolist()
pictures = sum(pictures, [])
pictures
```

Out[8]:
```
['990_left.jpg',
 '4496_left.jpg',
 '1113_left.jpg',
 '2432_left.jpg',
 '2123_left.jpg',
 '2204_left.jpg',
 '3120_left.jpg',
 '1038_left.jpg',
 '149_left.jpg',
 '2225_left.jpg',
 '2708_left.jpg',
 '448_left.jpg',
 '683_left.jpg',
 '2236_left.jpg',
 '3204_left.jpg',
 '2532_left.jpg',
 '384_left.jpg',
 '3394_left.jpg',
 '2209_left.jpg',
 '2832_left.jpg',
 '2135_left.jpg',
 '2507_left.jpg',
 '4219_left.jpg',
 '2855_left.jpg',
 '2932_left.jpg',
 '2101_left.jpg',
 '3153_left.jpg',
 '2262_left.jpg',
 '2145_left.jpg',
 '2216_left.jpg',
 '2846_left.jpg',
 '2171_left.jpg',
 '2157_left.jpg',
 '2140_left.jpg',
 '3420_left.jpg',
 '3163_left.jpg',
 '85_left.jpg',
 '3183_left.jpg',
 '2117_left.jpg',
 '2744_left.jpg',
 '2134_left.jpg',
 '2166_left.jpg',
 '2098_left.jpg',
 '2774_left.jpg',
 '1484_left.jpg',
 '2232_left.jpg',
 '907_left.jpg',
 '2543_left.jpg',
 '4531_left.jpg',
 '3311_left.jpg',
 '2470_left.jpg',
 '2185_left.jpg',
 '2153_left.jpg',
 '69_left.jpg',
 '2211_left.jpg',
 '2194_left.jpg',
 '2130_left.jpg',
 '2189_left.jpg',
 '2129_left.jpg',
 '2243_left.jpg',
 '3208_left.jpg',
 '2666_left.jpg',
 '999_left.jpg',
 '3172_left.jpg',
 '2452_left.jpg',
 '2872_left.jpg',
 '4209_left.jpg',
 '294_left.jpg',
 '2181_left.jpg',
 '3361_left.jpg',
 '2649_left.jpg',
 '3076_left.jpg',
 '2124_left.jpg',
 '3064_left.jpg',
 '2127_left.jpg',
 '2151_left.jpg',
 '103_left.jpg',
 '2095_left.jpg',
 '2461_left.jpg',
 '0_right.jpg',
 '2178_left.jpg',
 '2179_left.jpg',
 '809_left.jpg',
 '2692_left.jpg',
 '824_left.jpg',
 '2200_left.jpg',
 '1162_left.jpg',
 '477_left.jpg',
 '702_left.jpg',
 '2466_left.jpg',
 '2227_left.jpg',
 '2861_left.jpg',
 '1452_left.jpg',
 '2183_left.jpg',
 '160_left.jpg',
 '2158_left.jpg',
 '2180_left.jpg',
 '2207_left.jpg',
 '1415_left.jpg',
 '333_left.jpg',
 '1166_left.jpg',
 '2152_left.jpg',
 '2529_left.jpg',
 '3379_left.jpg',
 '2184_left.jpg',
 '877_left.jpg',
 '2849_left.jpg',
 '2176_left.jpg',
 '599_left.jpg',
 '2575_left.jpg',
 '2965_left.jpg',
 '2233_left.jpg',
 '2221_left.jpg',
 '1164_left.jpg',
 '2819_left.jpg',
 '2369_left.jpg',
 '691_left.jpg',
 '2150_left.jpg',
 '750_left.jpg',
 '695_left.jpg',
 '3057_left.jpg',
 '3025_left.jpg',
 '3112_left.jpg',
 '1444_left.jpg',
 '2126_left.jpg',
 '2146_left.jpg',
 '2247_left.jpg',
 '3423_left.jpg',
 '2136_left.jpg',
 '2377_left.jpg',
 '2195_left.jpg',
 '2099_left.jpg',
 '2346_left.jpg',
 '629_left.jpg',
 '2154_left.jpg',
 '2167_left.jpg',
 '2144_left.jpg',
 '2155_left.jpg',
 '2122_left.jpg',
 '2514_left.jpg',
 '4230_left.jpg',
 '2110_left.jpg',
 '2107_left.jpg',
 '4224_left.jpg',
 '1008_left.jpg',
 '3173_left.jpg',
 '2993_left.jpg',
 '2441_left.jpg',
 '347_left.jpg',
 '820_left.jpg',
 '2170_left.jpg',
 '2105_left.jpg',
 '2234_left.jpg',
 '2223_left.jpg',
 '1468_left.jpg',
 '4500_left.jpg',
 '2113_left.jpg',
 '2212_left.jpg',
 '553_left.jpg',
 '3102_left.jpg',
 '2480_left.jpg',
 '3128_left.jpg',
 '3121_left.jpg',
 '2722_left.jpg',
 '2375_left.jpg',
 '2876_left.jpg',
 '2448_left.jpg',
 '2121_left.jpg',
 '2163_left.jpg',
 '4577_left.jpg',
 '2215_left.jpg',
 '1102_left.jpg',
 '2434_left.jpg',
 '2138_left.jpg',
 '2175_left.jpg',
 '1200_left.jpg',
 '2206_left.jpg',
 '2235_left.jpg',
 '2169_left.jpg',
 '2931_left.jpg',
 '2446_left.jpg',
 '147_left.jpg',
 '2914_left.jpg',
 '2462_left.jpg',
 '2904_left.jpg',
 '2246_left.jpg',
 '2133_left.jpg',
 '3136_left.jpg',
 '3020_left.jpg',
 '3159_left.jpg',
 '945_left.jpg',
 '3067_left.jpg',
 '2182_left.jpg',
 '2385_left.jpg',
 '2116_left.jpg',
 '2128_left.jpg',
 '2210_left.jpg',
 '2890_left.jpg',
 '173_left.jpg',
 '2132_left.jpg',
 '191_left.jpg',
 '2219_left.jpg',
 '2239_left.jpg',
 '4309_left.jpg',
 '2111_left.jpg',
 '3401_left.jpg',
 '1987_left.jpg',
 '2108_left.jpg',
 '2399_left.jpg',
 '2192_left.jpg',
 '16_left.jpg',
 '2137_left.jpg',
 '1115_left.jpg',
 '3331_left.jpg',
 '2217_left.jpg',
 '2198_left.jpg',
 '439_left.jpg',
 '58_left.jpg',
 '1285_left.jpg',
 '2104_left.jpg',
 '4245_left.jpg',
 '2159_left.jpg',
 '2114_left.jpg',
 '2125_left.jpg',
 '2240_left.jpg',
 '2431_left.jpg',
 '3129_left.jpg',
 '611_left.jpg',
 '2482_left.jpg',
 '425_left.jpg',
 '2349_left.jpg',
 '4489_left.jpg',
 '2242_left.jpg',
 '330_left.jpg',
 '2149_left.jpg',
 '1435_left.jpg',
 '2778_left.jpg',
 '3249_left.jpg',
 '2147_left.jpg',
 '1483_left.jpg',
 '2168_left.jpg',
 '187_left.jpg',
 '2142_left.jpg',
 '2902_left.jpg',
 '1144_left.jpg',
 '3374_left.jpg',
 '119_left.jpg',
 '3075_left.jpg',
 '8_left.jpg',
 '2896_left.jpg',
 '560_left.jpg',
 '2609_left.jpg',
 '2156_left.jpg',
 '3382_left.jpg',
 '3424_left.jpg',
 '317_left.jpg',
 '2987_left.jpg',
 '2398_left.jpg',
 '2226_left.jpg',
 '726_left.jpg',
 '2199_left.jpg',
 '254_left.jpg',
 '3303_left.jpg',
 '2203_left.jpg',
 '3448_left.jpg',
 '2740_left.jpg',
 '2109_left.jpg',
 '2120_left.jpg',
 '2916_left.jpg',
 '2172_left.jpg']
```

In [9]:
```python
#sampling 30 photos
small_pics= random.sample(pictures, 270)
```

In [ ]:
```python

```

In [10]:
```python
#grab target if label matches
classifications=list()
for pic_name in small_pics:
    intermediate = theFile.loc[theFile['filename'] == pic_name]
    label = intermediate['Left-Diagnostic Keywords'].to_string().replace('\'','').replace('    ', '').r
eplace(',', '')
    label = ''.join([i for i in label if not i.isdigit()])
    classifications.append(label)
#classifications contains the "target" corresponding to small_pics
```

In [11]:
```python
classifications
```

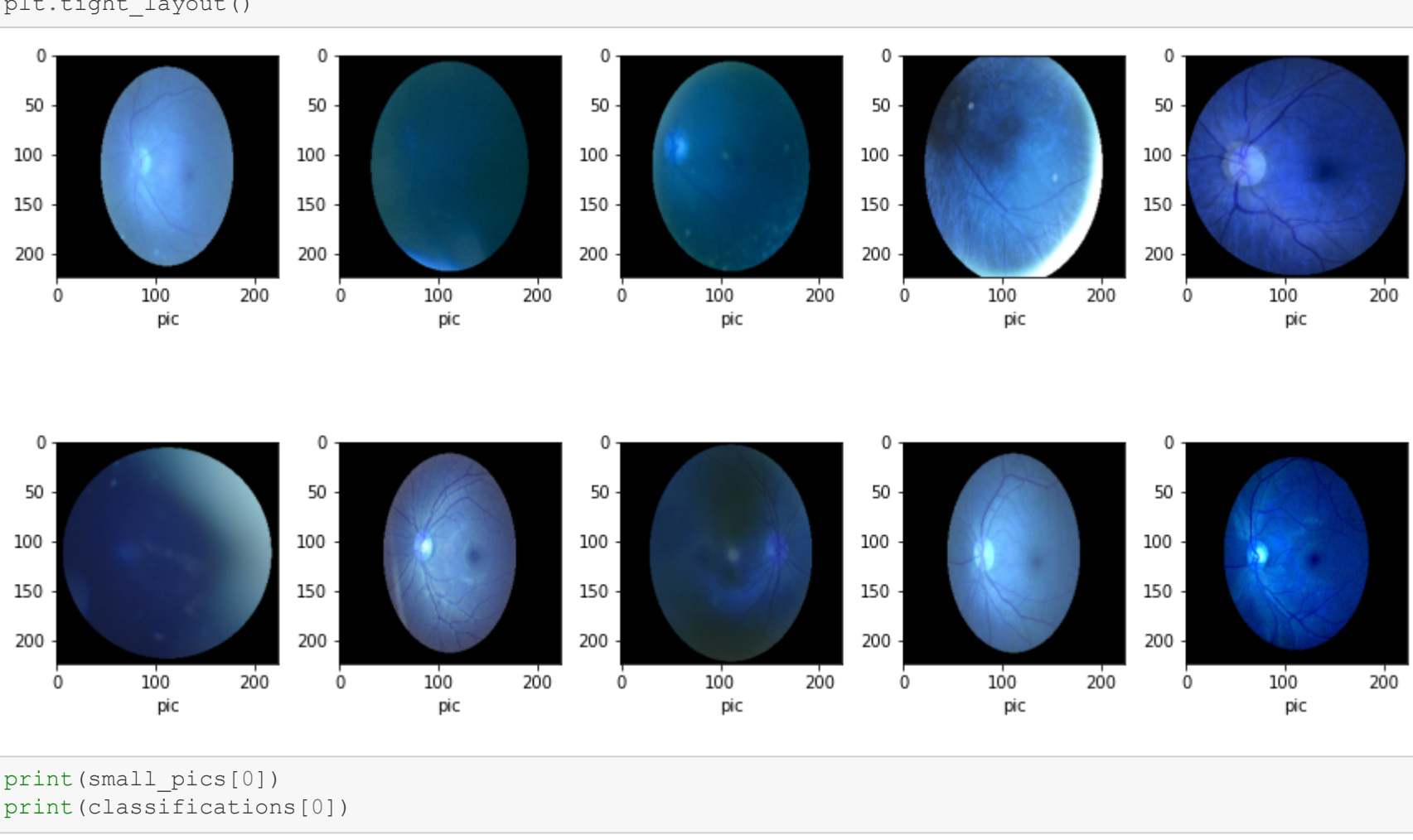
Out[11]:
```
['cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract',
 'cataract',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'normal fundus',
 'normal fundus',
 'normal fundus',
 'cataract',
 'cataract',
 'normal fundus',
 'cataract',
 'normal fundus',
 'cataract']
```

In [12]:
```python
import matplotlib.pyplot as plt
import cv2
actual_pics = []

img_res = 224

for file in small_pics:
    img = cv2.imread('../Challenge/images/' + file) #zero makes it greyscale
    #this is resizing all images
    img = cv2.resize(img, (img_res, img_res))
    actual_pics.append(img)
```

In [13]:
```python
actual_pics[0]
```

Out[13]:
```
array([[[0, 0, 0],
        [0, 0, 0],
        [0, 0, 0],
        ...,
        [0, 0, 0],
        [0, 0, 0],
        [0, 0, 0]],

       [[0, 0, 0],
        [0, 0, 0],
        [0, 0, 0],
        ...,
        [0, 0, 0],
        [0, 0, 0],
        [0, 0, 0]],

       [[0, 0, 0],
        [0, 0, 0],
        [0, 0, 0],
        ...,
        [0, 0, 0],
        [0, 0, 0],
        [0, 0, 0]],

       ...,

       [[0, 0, 0],
        [0, 0, 0],
        [0, 0, 0],
        ...,
        [0, 0, 0],
        [0, 0, 0],
        [0, 0, 0]],

       [[0, 0, 0],
        [0, 0, 0],
        [0, 0, 0],
        ...,
        [0, 0, 0],
        [0, 0, 0],
        [0, 0, 0]],

       [[0, 0, 0],
        [0, 0, 0],
        [0, 0, 0],
        ...,
        [0, 0, 0],
        [0, 0, 0],
        [0, 0, 0]]], dtype=uint8)
```

In [14]:
```python
plt.figure(figsize=(12,7))
for i in range(10):
    sample = random.choice(range(len(actual_pics)))
    image = actual_pics[sample]

    plt.subplot(2,5,i+1)
    plt.imshow(image)
    plt.xlabel('pic')
plt.tight_layout()
```



In [15]:
```python
print(small_pics[0])
print(classifications[0])
```

```
2204_left.jpg
cataract
```

In [16]:
```python
import numpy as np
from sklearn.model_selection import train_test_split
from tensorflow.keras import utils

#normalize our data !!
labels = {
    "cataract": 0,
    "normal fundus": 1
}

labels = np.array(labels)


for i in range(len(classifications)):
        if(classifications[i] == 'cataract'):
                classifications[i] = 0
        elif(classifications[i] == 'normal fundus'):
                classifications[i] = 1

classifications = utils.to_categorical(classifications)
```