Ćwiczenie 2

Celem ćwiczenia jest samodzielne wykorzystanie wiedzy na temat wczytywania i przetwarzania języku Python.

In [1]:
```
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import pandas as pd
import numpy as np
from sklearn.datasets import load_breast_cancer
from sklearn import preprocessing
data = load_breast_cancer()
from sklearn.feature_selection import SelectKBest
from sklearn.feature_selection import SelectPercentile, f_regression, mutual_inf
```

In [2]:
```
dir(data) # dir zwraca listę atrybutów danego obiektu
```

```
['DESCR', 'data', 'feature_names', 'filename', 'target', 'target_names']
```

Mając do dyspozycji bazę danych dotyczącą raka piersi:

- zwizualizuj znormalizowane dane w postaci wykresu 3D
- wybierz trzy cechy, które pozwalają na rozróżnienie od siebie przynajmniej dwóch klas
- użyj funkcji plot_surface do narysowania płaszczyny pozwalającej na rozróżnienie dwóch kl muszą być idealnie oddzielone)

In [3]:
```
X_data, y_data = data.data, data.target #podział danych na macierz cech i wektor
print (X_data.shape, y_data.shape)
print (X_data[0], y_data[0])
print (data.target_names)    #nazwy klas
print (data.feature_names)   #nazwy cech
#print (data.DESCR)
```

```
(569, 30) (569,)
[1.799e+01 1.038e+01 1.228e+02 1.001e+03 1.184e-01 2.776e-01 3.001e-01
 1.471e-01 2.419e-01 7.871e-02 1.095e+00 9.053e-01 8.589e+00 1.534e+02
 6.399e-03 4.904e-02 5.373e-02 1.587e-02 3.003e-02 6.193e-03 2.538e+01
 1.733e+01 1.846e+02 2.019e+03 1.622e-01 6.656e-01 7.119e-01 2.654e-01
 4.601e-01 1.189e-01] 0
['malignant' 'benign']
['mean radius' 'mean texture' 'mean perimeter' 'mean area'
 'mean smoothness' 'mean compactness' 'mean concavity'
 'mean concave points' 'mean symmetry' 'mean fractal dimension'
 'radius error' 'texture error' 'perimeter error' 'area error'
 'smoothness error' 'compactness error' 'concavity error'
 'concave points error' 'symmetry error' 'fractal dimension error'
 'worst radius' 'worst texture' 'worst perimeter' 'worst area'
 'worst smoothness' 'worst compactness' 'worst concavity'
 'worst concave points' 'worst symmetry' 'worst fractal dimension']
```

In [4]:
```
skala = preprocessing.MinMaxScaler()
pandy = pd.DataFrame(data = skala.fit_transform(data.data),
                                index = data.target, columns = data.feature_names)
```

In [5]:
```
pandy
```

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	f dime
0	0.521037	0.022658	0.545989	0.363733	0.593753	0.792037	0.703140	0.731113	0.686364	0.605
0	0.643144	0.272574	0.615783	0.501591	0.289880	0.181768	0.203608	0.348757	0.379798	0.141
0	0.601496	0.390260	0.595743	0.449417	0.514309	0.431017	0.462512	0.635686	0.509596	0.211
0	0.210090	0.360839	0.233501	0.102906	0.811321	0.811361	0.565604	0.522863	0.776263	1.000
0	0.629893	0.156578	0.630986	0.489290	0.430351	0.347893	0.463918	0.518390	0.378283	0.186
...	...	...	...	...	...	...	...	...	...	...
0	0.690000	0.428813	0.678668	0.566490	0.526948	0.296055	0.571462	0.690358	0.336364	0.132
0	0.622320	0.626987	0.604036	0.474019	0.407782	0.257714	0.337395	0.486630	0.349495	0.113
0	0.455251	0.621238	0.445788	0.303118	0.288165	0.254340	0.216753	0.263519	0.267677	0.137
0	0.644564	0.663510	0.665538	0.475716	0.588336	0.790197	0.823336	0.755467	0.675253	0.425
1	0.036869	0.501522	0.028540	0.015907	0.000000	0.074351	0.000000	0.000000	0.266162	0.187

569 rows × 30 columns

In [6]:
```
Bestest = SelectKBest(mutual_info_regression, k=3)
Bestest.fit(pandy, data.target)
columns = Bestest.get_support(indices=True)
```

In [7]:
```
1 Bestpandy = pandy.iloc[:,columns]
2 Bestpandy
```

	worst radius	worst perimeter	worst area
0	0.620776	0.668310	0.450698
0	0.606901	0.539818	0.435214
0	0.556386	0.508442	0.374508
0	0.248310	0.241347	0.094008
0	0.519744	0.506948	0.341575
...	...	...	...
0	0.623266	0.576174	0.452664
0	0.560655	0.520892	0.379915
0	0.393099	0.379949	0.230731
0	0.633582	0.668310	0.402035
1	0.054287	0.043578	0.020497

569 rows × 3 columns

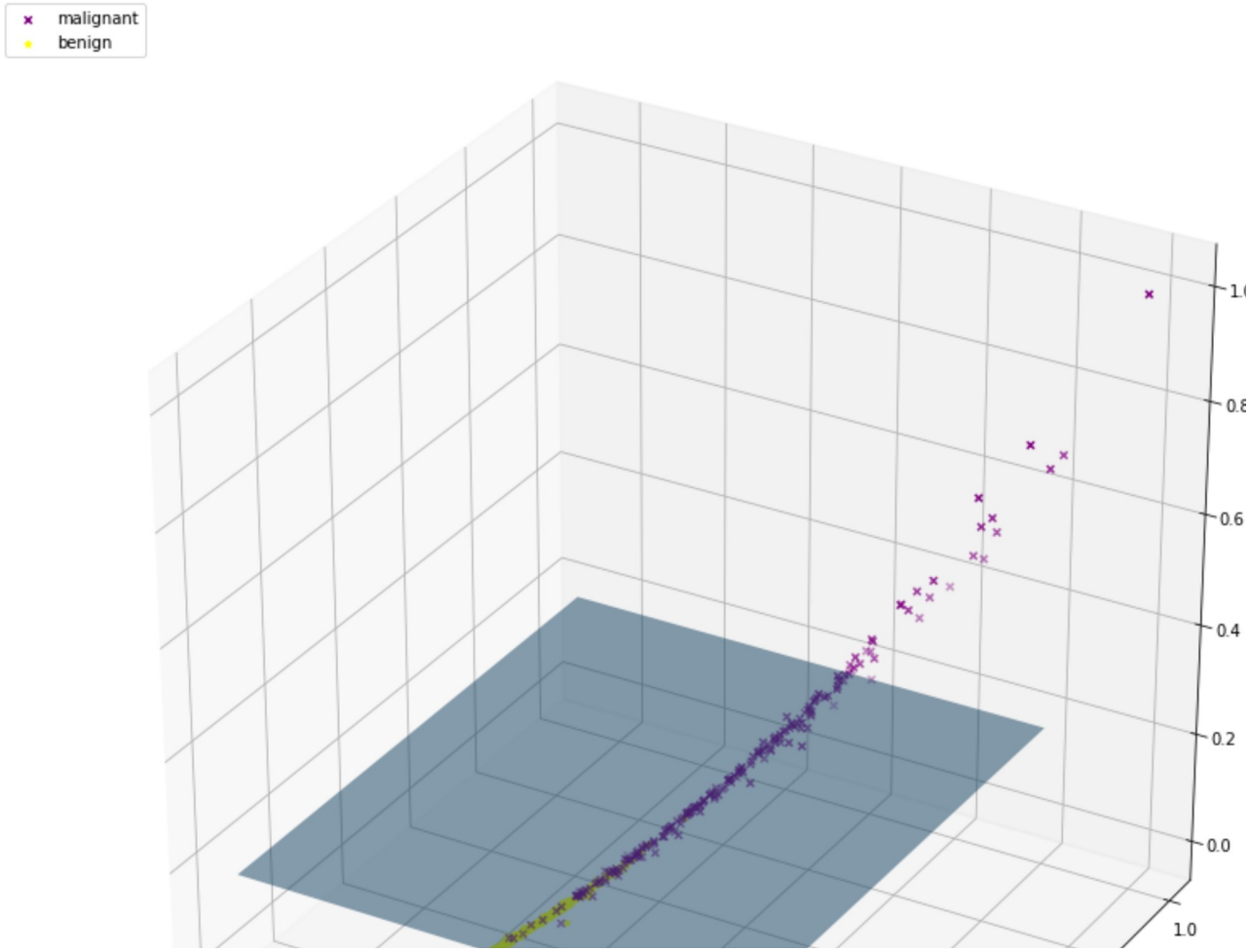
In [8]:
```
1 X1, X2, Y1, Y2, Z1, Z2 = Bestpandy['worst radius'][0], Bestpandy['worst radius']
2                          Bestpandy['worst perimeter'][0], Bestpandy['worst perim
3                          Bestpandy['worst area'][0], Bestpandy['worst area'][1]
```

In [9]:
```
1 dist = [[(Z1.median()+Z2.median())/2]*2]*2
```

In [ ]:
```
1
```

In [10]:

```
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D

fig = plt.figure(figsize=(15, 15))

ax = fig.add_subplot(111, projection='3d')
ax.scatter(X1, Y1, Z1, c='purple', marker='x', label=data.target_names[0]);
ax.scatter(X2, Y2, Z2, c='yellow', marker='*', label=data.target_names[1]);
ax.plot_surface(np.array([[-.3, -.3], [.8, .8]]),
                    np.array([[0, 1], [0, 1]]),
                    np.array(dist), alpha=.5);


ax.set_xlabel(Bestpandy.columns[0])
ax.set_ylabel(Bestpandy.columns[1])
ax.set_zlabel(Bestpandy.columns[2])

plt.legend(loc='upper left')

plt.show()
```



In [ ]:
```
1
```

In [ ]:
```
1
```

In [11]:
```
1 %matplotlib inline
```

Wczytywanie i przetwarzanie danych w postaci c z wykorzystaniem formatu csv

---

Ćwiczenie 3

Celem ćwiczenia jest:

- wyświetlanie macierzy jako obrazu;
- zapoznanie z formatem csv;
- przećwiczenie składni Pythona podczas implementacji algorytmu wczytywania bazy danych

---

Pobierz i rozpakuj bazę danych GTSRB
http://benchmark.ini.rub.de/?section=gtsrb&subsection=dataset#Downloads (http://benchmark.ir /?section=gtsrb&subsection=dataset#Downloads)
Wczytywanie bazy danych z pliku:

In [12]:
```
1 from imageio import imread
2 file_path = r"C:\Users\sticz\Desktop\Magisterka sezon drugi\Podstawy Uczenia Mas
3 img = imread(file_path)
4 img.shape
```

```
(30, 29, 3)
```

In [13]:
```
1 img.size
```

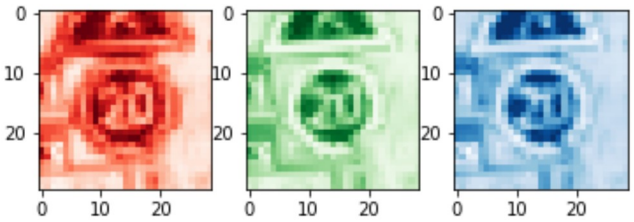
```
2610
```

In [14]:
```
1 import sklearn as sk
2 import matplotlib.pyplot as plt
```

Jaki format ma jeden rekord? Jest to obraz w formacie .ppm. Jakiego rozmiaru jest to obrazek?

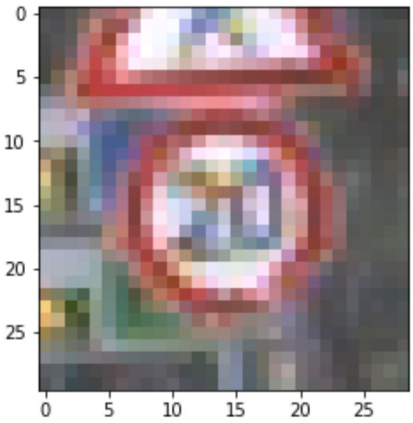
Wyświetlamy wszystkie kanały osobno za pomoca matplotlib.

In [15]:
```
f, (ax1, ax2, ax3) = plt.subplots(1, 3)
ax1.imshow(img[:,:,0],cmap='Reds')
ax2.imshow(img[:,:,1],cmap='Greens')
ax3.imshow(img[:,:,2],cmap='Blues')
plt.show()
```



Wyświetlamy wszystkie kanały na raz

In [16]:
```
plt.figure()
plt.imshow(img)
plt.show()
```



Korzystając z pliku csv wczytaj obrazy z bazy danych.
Następnie wyświetl 25 losowych obrazów na jednej kanwie, np. używając subplots.

In [17]:
```
import pandas
pandy = pandas.read_csv(r"C:\Users\sticz\Desktop\Magisterka sezon drugi\Podstawy
```

In [18]:
```
1 pandy
```

	Filename	Width	Height	Roi.X1	Roi.Y1	Roi.X2	Roi.Y2	ClassId
0	00000_00000.ppm	29	30	5	6	24	25	0
1	00000_00001.ppm	30	30	5	5	25	25	0
2	00000_00002.ppm	30	30	5	5	25	25	0
3	00000_00003.ppm	31	31	5	5	26	26	0
4	00000_00004.ppm	30	32	5	6	25	26	0
...	...	...	...	...	...	...	...	...
205	00006_00025.ppm	85	87	8	7	77	80	0
206	00006_00026.ppm	92	95	8	8	83	87	0
207	00006_00027.ppm	97	100	9	8	89	92	0
208	00006_00028.ppm	105	109	9	10	95	100	0
209	00006_00029.ppm	112	118	10	11	103	108	0

210 rows × 8 columns

In [19]:
```
1 pandy_int = pandy[['Filename','ClassId']]
2 display(pandy_int)
```

	Filename	ClassId
0	00000_00000.ppm	0
1	00000_00001.ppm	0
2	00000_00002.ppm	0
3	00000_00003.ppm	0
4	00000_00004.ppm	0
...	...	...
205	00006_00025.ppm	0
206	00006_00026.ppm	0
207	00006_00027.ppm	0
208	00006_00028.ppm	0
209	00006_00029.ppm	0

210 rows × 2 columns

In [20]:

```
1 pandy_25 = pandy_int.sample(25)
2 pandy_25
```

	Filename	ClassId
185	00006_00005.ppm	0
108	00003_00018.ppm	0
151	00005_00001.ppm	0
195	00006_00015.ppm	0
48	00001_00018.ppm	0
193	00006_00013.ppm	0
79	00002_00019.ppm	0
125	00004_00005.ppm	0
65	00002_00005.ppm	0
182	00006_00002.ppm	0
118	00003_00028.ppm	0
131	00004_00011.ppm	0
84	00002_00024.ppm	0
38	00001_00008.ppm	0
0	00000_00000.ppm	0
165	00005_00015.ppm	0
11	00000_00011.ppm	0
39	00001_00009.ppm	0
26	00000_00026.ppm	0
8	00000_00008.ppm	0
128	00004_00008.ppm	0
73	00002_00013.ppm	0
5	00000_00005.ppm	0
179	00005_00029.ppm	0
111	00003_00021.ppm	0

In [21]:
```
1 filename = pandy_25[['Filename']]
2 display(filename)
```

	Filename
185	00006_00005.ppm
108	00003_00018.ppm
151	00005_00001.ppm
195	00006_00015.ppm
48	00001_00018.ppm
193	00006_00013.ppm
79	00002_00019.ppm
125	00004_00005.ppm
65	00002_00005.ppm
182	00006_00002.ppm
118	00003_00028.ppm
131	00004_00011.ppm
84	00002_00024.ppm
38	00001_00008.ppm
0	00000_00000.ppm
165	00005_00015.ppm
11	00000_00011.ppm
39	00001_00009.ppm
26	00000_00026.ppm
8	00000_00008.ppm
128	00004_00008.ppm
73	00002_00013.ppm
5	00000_00005.ppm
179	00005_00029.ppm
111	00003_00021.ppm

In [22]:
```
1 filenames = filename.values.tolist()
```

In [23]:
```
1 filenames
```

```
[['00006_00005.ppm'],
 ['00003_00018.ppm'],
 ['00005_00001.ppm'],
 ['00006_00015.ppm'],
 ['00001_00018.ppm'],
 ['00006_00013.ppm'],
 ['00002_00019.ppm'],
 ['00004_00005.ppm'],
 ['00002_00005.ppm'],
 ['00006_00002.ppm'],
 ['00003_00028.ppm'],
 ['00004_00011.ppm'],
 ['00002_00024.ppm'],
 ['00001_00008.ppm'],
 ['00000_00000.ppm'],
 ['00005_00015.ppm'],
 ['00000_00011.ppm'],
 ['00001_00009.ppm'],
 ['00000_00026.ppm'],
 ['00000_00008.ppm'],
 ['00004_00008.ppm'],
 ['00002_00013.ppm'],
 ['00000_00005.ppm'],
 ['00005_00029.ppm'],
 ['00003_00021.ppm']]
```

In [24]:
```
1 import keras
2 from keras.preprocessing.image import ImageDataGenerator, load_img
```

```
Using TensorFlow backend.
```

In [25]:
```
1 filenames_full = r"C:\Users\sticz\Desktop\Magisterka sezon drugi\Podstawy Uczeni
```

In [26]:
```
1 filenames_full
```

	Filename
185	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
108	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
151	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
195	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
48	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
193	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
79	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
125	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
65	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
182	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
118	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
131	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
84	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
38	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
0	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
165	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
11	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
39	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
26	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
8	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
128	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
73	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
5	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
179	C:\Users\sticz\Desktop\Magisterka sezon drugi\...
111	C:\Users\sticz\Desktop\Magisterka sezon drugi\...

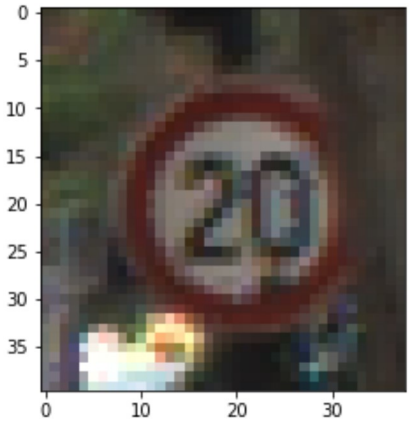
In [27]:
```
filenamesfull_list = filenames_full.values.tolist()
filenamesfull_list
```

```
[['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00006_00005.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00003_00018.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00005_00001.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00006_00015.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00001_00018.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00006_00013.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00002_00019.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00004_00005.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00002_00005.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00006_00002.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00003_00028.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00004_00011.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00002_00024.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00001_00008.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00000_00000.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00005_00015.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00000_00011.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00001_00009.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00000_00026.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00000_00008.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00004_00008.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00002_00013.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00000_00005.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00005_00029.ppm'],
 ['C:\\Users\\sticz\\Desktop\\Magisterka sezon drugi\\Podstawy Uczenia Maszynowego\\PUM-master\\Laboratoriur
ining\\Images\\00000\\\\00003_00021.ppm']]
```

In [28]:
```
from pathlib import Path
```

In [29]:

```
1 img = imread(Path(filenamesfull_list[0][0]))
2 plt.figure()
3 plt.imshow(img)
4 plt.show()
```



In [30]:

```
1 IMAGE_WIDTH=128
2 IMAGE_HEIGHT=128
3 IMAGE_SIZE=(IMAGE_WIDTH, IMAGE_HEIGHT)
```

In [31]:
```
1 pandy_25
```

	Filename	ClassId
185	00006_00005.ppm	0
108	00003_00018.ppm	0
151	00005_00001.ppm	0
195	00006_00015.ppm	0
48	00001_00018.ppm	0
193	00006_00013.ppm	0
79	00002_00019.ppm	0
125	00004_00005.ppm	0
65	00002_00005.ppm	0
182	00006_00002.ppm	0
118	00003_00028.ppm	0
131	00004_00011.ppm	0
84	00002_00024.ppm	0
38	00001_00008.ppm	0
0	00000_00000.ppm	0
165	00005_00015.ppm	0
11	00000_00011.ppm	0
39	00001_00009.ppm	0
26	00000_00026.ppm	0
8	00000_00008.ppm	0
128	00004_00008.ppm	0
73	00002_00013.ppm	0
5	00000_00005.ppm	0
179	00005_00029.ppm	0
111	00003_00021.ppm	0

In [32]:

```
sample_test = pandy_25.sample(25)
sample_test.head(25)
plt.figure(figsize=(50, 48))
a = 0
lista = []
for index, row in sample_test.iterrows():
    index = a
    filename = row['Filename']
    category = row['ClassId']
    img = load_img(filenamesfull_list[a][0], target_size=IMAGE_SIZE)
    plt.subplot(10, 5, index+1)
    plt.imshow(img)
    plt.xlabel(filename + '(' + "{}".format(category) + ')' ) #filename +
    a = a+1
    lista.append(category)
plt.tight_layout()
plt.show()
```

Type *Markdown* and LaTeX: α^2

In [ ]:

```
1
```