before_pass_module1

August 29, 2024

```
[ ]: !pwd
     # %matplotlib inline
     # %pip install pickleshare
```

/home/haipn/Clone/Research_ThyroidFNA_ClassAI/phase2_280824/notebooks/explore

```
[ ]: %cd ../..
```

/home/haipn/Clone/Research_ThyroidFNA_ClassAI/phase2_280824

```
[ ]: from src.data.explorer.before_pass_module1 import *
```

```
[ ]: !ls data
```

```
1724896178_A_set  1724896178_C_set  Origin_Bx-July.zip
1724896178_B_set  1724896178_D_set  raw
```

```
[ ]: folder_path = 'data/'
     images = load_random_images(folder_path=folder_path, num_images=100)
     print(len(images), len(images[0]), len(images[1]))
     print(images[0][:5], images[1][:5], sep='\n')
```

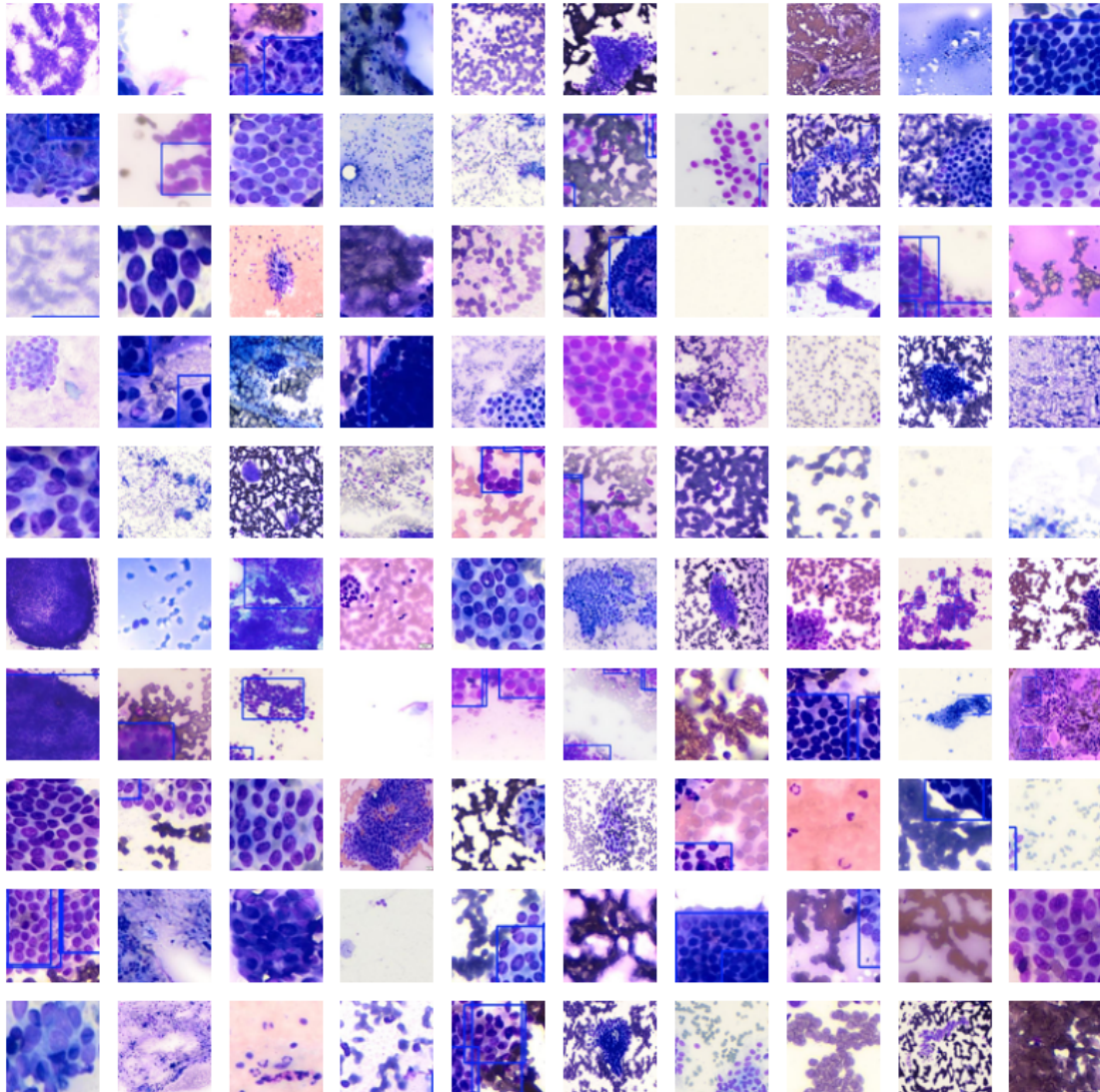```
Số lượng ảnh trong mỗi folder:
data/1724896178_C_set/train/B6: 4368 ảnh
data/1724896178_C_set/train/B2: 2696 ảnh
data/1724896178_C_set/train/B5: 3024 ảnh
data/1724896178_C_set/valid/B6: 936 ảnh
data/1724896178_C_set/valid/B2: 576 ảnh
data/1724896178_C_set/valid/B5: 648 ảnh
data/raw/B6: 781 ảnh
data/raw/B2: 482 ảnh
data/raw/B5: 541 ảnh
data/1724896178_A_set/train/B6: 546 ảnh
data/1724896178_A_set/train/B2: 337 ảnh
data/1724896178_A_set/train/B5: 378 ảnh
data/1724896178_A_set/test/B6: 118 ảnh
data/1724896178_A_set/test/B2: 73 ảnh
data/1724896178_A_set/test/B5: 82 ảnh
data/1724896178_A_set/valid/B6: 117 ảnh
data/1724896178_A_set/valid/B2: 72 ảnh
```

```
data/1724896178_A_set/valid/B5: 81 ảnh
data/1724896178_D_set/train/B6: 6552 ảnh
data/1724896178_D_set/train/B2: 4044 ảnh
data/1724896178_D_set/train/B5: 4536 ảnh
data/1724896178_D_set/valid/B6: 1404 ảnh
data/1724896178_D_set/valid/B2: 864 ảnh
data/1724896178_D_set/valid/B5: 972 ảnh
data/1724896178_B_set/train/B6: 546 ảnh
data/1724896178_B_set/train/B2: 337 ảnh
data/1724896178_B_set/train/B5: 378 ảnh
data/1724896178_B_set/valid/B6: 117 ảnh
data/1724896178_B_set/valid/B2: 72 ảnh
data/1724896178_B_set/valid/B5: 81 ảnh
2 100 100
[<PIL.Image.Image image mode=RGB size=64x64 at 0x7F46776CDA30>, <PIL.Image.Image
image mode=RGB size=64x64 at 0x7F4675727620>, <PIL.Image.Image image mode=RGB
size=64x64 at 0x7F4676B1CB60>, <PIL.Image.Image image mode=RGB size=64x64 at
0x7F46F0B7BD10>, <PIL.Image.Image image mode=RGB size=64x64 at 0x7F46F0B7BB30>]
('B6', 'B6', 'B6', 'B2', 'B2')
```

[ ]: ```python
plot_images(images=images[0])
```

```
reduced_data = reduce_dimensionality(images=images[0])
```

```
plot_interactive_3d(data=reduced_data, labels=images[1])
```