CS272 - Computer Vision II Quiz

Name:

Student ID:

1. (5 points) What is the output of the following code? (_____)

```
x = 7
y = 2
print(x % y)
print(x // y)
print(x ** y)
```

1
3
49

2. (5 points) What is the output of the following code? (_____)

```
t = 'a b c d e'
for i in range(10, 0, -2):
    for j, c in enumerate(t):
        if i == j:
            print(c)
            break
```

e
d
c
b

3. (5 points) What is the output of the following code? (_____)

```
def func(num):
    num += 2
    return num

func(2)
print(num)
```

NameError

4. (5 points) What is the output of the following code? (_____)

```
dict = defaultdict(int)
print(dict['name'])
```

0

5. (5 points) What is the output of the following code? (_____)

```
import numpy as np

a = np.eye(3, dtype=float)
b = np.array(range(9)).reshape(-1, 3)
a = a + b
print(a[1][2])
```

5.0

6. (5 points) What is the output of the following code? (_____)

```
import torch

a = torch.tensor(range(9)).view(-1, 3)
b = a > -1
c = torch.ones_like(a)
print(b is c)
```

False

7. (20 points) What is the value of a, b, c, d? (_____)

```
import torch

m = torch.tensor([[1, 2], [0, 0]]).T
n = torch.ones_like(m)
p = torch.tensor([1, 2])
q = torch.tensor([[1, 2]]).T
a = m * n
b = m + p
c = torch.cat((m, q), 1)
d = a.sum(0)
```

a [[1, 0], [2, 0]]

b [[2, 2], [3, 2]]

c [[1, 0, 1], [2, 0, 2]]

d [3, 0]

8. (20 points) Complete the following simple CNN for 3 * 32 * 32 images?

```
import torch
import torch.nn as nn
import torch.nn.functional as F

class TestNet(nn.Module):
    def __init__(self, num_classes, in_channels):
        super().__init__()
        self.conv1 = nn.Conv2d(in_channels=in_channels,
        out_channels=64, kernel_size=3, stride=1, padding=1)
        self.pool = nn.MaxPool2d(kernel_size=2, stride=2)
        self.conv2 = nn.Conv2d(in_channels=64,
        out_channels=128, kernel_size=5, stride=1, padding=0)
        self.fc1 = nn.Linear(in_features=6x6x128, out_features=128)
        self.fc2 = ____ \ 1 ____ / 128 ____ \ 1 ____ num_classes

    def forward(self, x):
        x = self.pool(F.relu(self.conv1(x)))
        x = self.pool(F.relu(self.conv2(x)))
        x = x.flatten()
        x = F.relu(self.fc1(x))
        x = self.fc2(x)
        return x

net = TestNet(10, 3)
```

9. (15 points) What's the relationship/difference between eye fixation detection and salient object detection?

They are both visual saliency detection methods. As an early work, eye fixation detection pays more attention to the mechanism of human visual attention, and predicts the possibility of human eyes staying at a certain position in the scene. Salient object detection emphasizes the accurate prediction of the whole salient target in the scene and the acquisition of clear salient object boundaries, which provides more direct and effective information for object level visual tasks.

10. (15 points) How to get the ground-truth for salient object detection and eye fixation detection?

Salient object detection: Arrange a certain number of representative testees to mark prominent objects by object bounding box calibration or pixel level calibration. The number of significant objects can be limited or unlimited. Then average multiple testees' results to obtain ground truth.

Eye fixation detection: Arrange a certain number of representative testees, and accurately record the eye movement data of each testee in the free viewing image mode by the EMR, including gaze trajectory and gaze position. Reasonably set the viewing time of each image. Average the eye movement data of each testee, and carry out two-dimensional Gaussian convolution for all fixation positions to obtain a continuous fixation density map, and then obtain the ground truth.