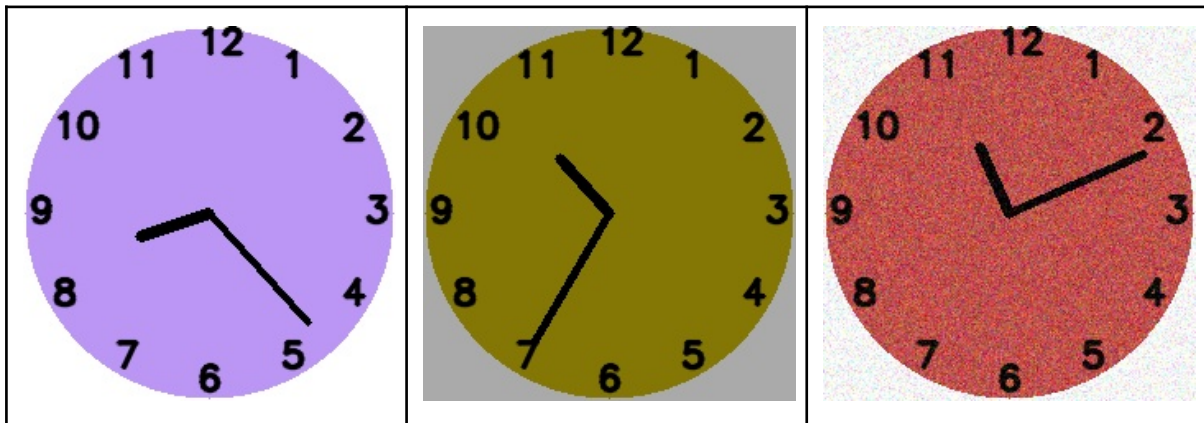


## AI51601/CSE54501 Computer Vision Assignment 1 [100 points]

**Deadline : until Nov. 22, 23:59 until Nov. 29, 23:59**

- If you have any questions, please post it in the blackboard/discussion/'assignment 1' board.
- Please use the PyTorch library when implementing CNN network. Other deep learning libraries (eg. Tensorflow, Keras) are not allowed.

1. [30 points] Using the OpenCV library, please write a code (Prob1.py) for drawing clock images given two arguments (ie. hour, minute) and meeting the conditions, as follows:



### [Conditions]

- The image size is 227 pixels x 227 pixels x 3 channels.
- The clock is a circle and it has a radius of 112 pixels.
- The clock and background have random colors (but they are always different).
- The clock and background images can have random noises as the 3rd figure.
- The minute and hour hands start from the center of the image.
- The black-colored hour hand length is 45 pixels and the thickness is 5.
- The black-colored minute hand length is 90 pixels and the thickness is 3.
- Digits of 1~12 in black color are located as the clock and having fonts as follows:

```
font = cv2.FONT_HERSHEY_SIMPLEX
font_scale = 0.7
font_thickness = 2
cv2.putText(.. , .. , .. , font, font_scale, (0, 0, 0), font_thickness,
cv2.LINE_AA)
```

Example image are uploaded here:

[https://drive.google.com/drive/folders/1\\_zdugouxC2hSaLLsv\\_iCd5pbKHSSShhr1?usp=sharing](https://drive.google.com/drive/folders/1_zdugouxC2hSaLLsv_iCd5pbKHSSShhr1?usp=sharing)

2. [20 points] Please write a code (Prob2.py) that implements the CNN network predicting the hour and the minute from the 227x227x3 sized images that we obtained from Problem 1.

- You need to use the CNN network to predict the values. You can formulate it as either classification or regression task. Or you can perform both, by adding two losses (ie. multi-task learning).
- You are allowed to implement the code to load and manipulate the already pre-trained network (e.g. ImageNet pre-trained ResNet).

[Tips] You can decide the output of the CNN network not being the raw hour and minute values. Rather, you can formulate it to output the degrees of hour and minute hands. Or, you can estimate their x and y coordinate values or heat maps instead.

[Tips] In a regression task, the output space is too wide when you try to estimate the values from  $-\infty$  to  $\infty$ . Normally, we normalize the ground-truth/output values to be confined in 0 and 1 ranges. Also, you can use the sigmoid layer as the last layer to output values within 0 and 1 range and train the network using the MSE loss. Subsequently, you can un-normalize the output values to be the original values.

3. [10 points] Please implement a code (Prob3.py) for training the CNN network using the data you created in Problem 1 and train it. You can generate a sufficient amount of training data to properly train your network. You need to save the learned weights of CNN network and submit it in the submission file in the form of .pth file.
4. [10 points] Please implement a code (Prob4.py) for testing the CNN network for an image. The code inputs an image and it has to output the proper hour and minute.
5. [15 points] Please analyze the successful and failure cases of your CNN network. How accurately it can detect the hour and the minute from the images? In our preliminary implementation, we noticed that deep networks can accurately infer hours and minutes. We will measure the accuracy of your networks using 100 random samples, allowing 5 minutes as the error margin: for example we will regard it as correct if you estimate 10h 40m from 10h 38m images. The accuracy of the model will count for 15 points. But, if the accuracy is too low, some points will be also deducted in problems 2-3, as it is the indication of non-effective design of the model.
6. [15 points] Please note and apply the below details.
- a. Send all files(.zip) via **Blackboard**.
  - b. In **20231111\_SeungryulBaek\_cv\_ass1.zip** file, you need to have four python files (\*.py), a pre-trained network weight (.pth) and a report as follows:

<Example>

20151111\_SeungryulBaek\_cv\_ass1.zip

```
--Prob1.py
--Prob2.py
--Prob3.py
--Prob3.pth
--Prob4.py
--Report.pdf
```

- c. You need to write a report within 10 pages, properly summarizing your solutions or results for each problem.
- d. Code should have some **comment** that increase readability(it is also grading points)
- e. Your pre-trained weights in .pth need to be properly loaded.