Report Lab3

Task 1a :

The first task consists in using the AlexNet model to classify skin images. The model is made of convolution layers and three dense layers. The parameters are the following : LR=0.0001, batch_size=8, base_dense=8, optimizer=Adam, loss_function=binary_accuracy and the activation function is relu. We use only n_epochs=50.
So the result is not surprising the loss curves for the train set **do not converge** due to the small number of epochs. Besides, the loss curve for the validation set is **overfitting**. However, the accuracy is not so bad for both sets it reaches roughly 0.85.
Now with **dropout layers** added, there is **no overfitting** but the results are worse. Indeed the loss values are bigger and the accuracy for the test set reaches 0.80 and for the train set only 0.65.
Noise is added because of the dropout layers.

Task 1b :

Now if we add **batch normalization** at each convolutional layer we hope to see better results.
When we have batch normalization and dropout layers, it seems that the model learns faster but the loss curves do not converge due to the small number of epochs.
The final accuracy value of training accuracy is 0.7 more than in task 1a. At 20 epochs the training accuracy value is the same as in the previous task.
The effect of batch normalization is normalizing all features before the activation function. It **makes the learning more faster** but it is not a good idea to use it only to do the regularization job as we can see if we use only the batch normalization alone.

Indeed when we use only batch normalization, the model is overfitting and so even if we have the 1 accuracy value for the test set, we can't rely on this model.

Task 1c :

Now we change some hyperparameters as the learning rate LR=1e-5 and the number of epochs to 80. We try the model with and without batch normalization.
Without batch normalization the validation accuracy reaches 0.80. We can observe that the loss values are quite high and that there is noise.
With batch normalization, the validation accuracy is now 0.85 and the loss values are lower.
In both cases, the validation accuracy is much better than the train accuracy.
Once again the effect of batch normalization is to boost the speed for learning task.

Task 1d :

Now we repeat task 1c with 150 epochs and we observe that the results are similar. With batch normalization, the validation and test accuracy are better than without. The loss curves converge more also. Batch normalization is good for generalization because the results are better with new data.

<u>Task 2a and b</u> :

We now try to use spatial dropout instead of batch normalization. In task 2a we use 150 epochs and then in tak 2b we use 250 epochs. With spatial dropout, the loss curves do not converge and the validation and train accuracies are smaller. Spatial dropout seems not to be a good solution.

<u>Task 4 and 5</u> :

The goal of this task is to use data augmentation to overcome the problem of small dataset. With AlexNet model applied on skin images is rather good. The loss values do not go under 0.5. The validation accuracy is still 0.80.
It is similar with VGG16 model. Besides the dropout layers add a lot of noise.
But when using VGG16 on bone images the result is better. The loss value is approximately 0.2 and the validation accuracy is more than 0.9 and the train accuracy is 0.7.
The data augmentation is use to have more data and so the model is better trained.

<u>Task 7</u> :

Now we want to try transfer learning, use a pre-trained model convolutional layers to extract features and use them in another small MLP model. We use a VGG16 pre trained model and then a little MLP model to perform the classification task.
With transfer learning the results for skin images and bone images are really good.
The accuracy values are really high and the convergence is fast.

<u>Task 8</u> :

We want to see which regions the model focuses on to do the classification task. To do this we use a heatmap. What can be observed is that even if the accuracy from the Bone images is really good, the classification seems not to be according to the fracture zone. This is problematic.