1)

Filter the noise: make a convolution with a Gaussian kernel.
Take a derivative: make a convolution with. [-1 0 1]

We can combine both of them.

Two issues corresponding to the peak of the magnitude of the convolved image:

- Should be a local maximum
- It should be sufficiently high.

3)
A.
The overfitting problem is caused when the model fits the dataset too much when the training error is low, but the testing error increases. This might be caused by running the training process with a higher number of epochs.

Modifications

- Shuffling the dataset of training and testing
- Decrease the number of input features
- Use cross-validation techniques

B.
Batch Gradient Descent: The batch gradient descent evaluates all of the training instances and determines the error for each example in the training dataset before updating the model. Every epoch ends with a model update performed by the batch gradient. For large datasets, model updating and training become slow.

Mini-batch GD: Splits the training dataset into small batches and then calculates model error and update coefficient. This algorithm updates the model very frequently, which allows more robust convergence, unlike batch gradient descent.