Image classification

Data Description:You are provided with a dataset of images of plant seedlings at various stages of grown. Each image has a filename that is its  
unique id. The dataset comprises 12 plant species. The goal of the project is to create a classifier capable of determining a plant's species from a photo.

**Context:**Can you differentiate a weed from a crop seedling?  
The ability to do so effectively can mean better crop yields and better stewardship of the environment.  
The Aarhus University Signal Processing group, in collaboration with University of Southern Denmark, has  
recently released a dataset containing images of unique plants belonging to 12 species at several growth stages  
**Objective:**The goal of the project is to create a classifier capable of determining a plant's species from a photo.

**Learning Outcomes:**• Pre-processing of image data.  
• Visualization of images.  
• Building CNN.  
• Evaluate the Model.  
• The motive of the project is to make the learners capable to handle images/image classification problems, during this  
process you should also be capable to handle real image files, not just limited to a numpy array of image pixels.

**Steps and tasks:**1. Import the libraries, load dataset, print shape of data, visualize the images in dataset.   
2. Data Pre-processing:   
a. Normalization.  
b. Gaussian Blurring.  
c. Visualize data after pre-processing.  
3. Make data compatible:   
a. Convert labels to one-hot-vectors.  
b. Print the label for y\_train[0].  
c. Split the dataset into training, testing, and validation set.  
(Hint: First split images and labels into training and testing set with test\_size = 0.3. Then further split test data  
into test and validation set with test\_size = 0.5)  
d. Check the shape of data, Reshape data into shapes compatible with Keras models if it’s not already. If it’s  
already in the compatible shape, then comment in the notebook that it’s already in compatible shape.  
4. Building CNN:   
a. Define layers.  
b. Set optimizer and loss function. (Use Adam optimizer and categorical crossentropy.)  
5. Fit and evaluate model and print confusion matrix. (10 Marks)  
6. Visualize predictions for x\_test[2], x\_test[3], x\_test[33], x\_test[36], x\_test[59].