Step 1 – Generate you feature vectors

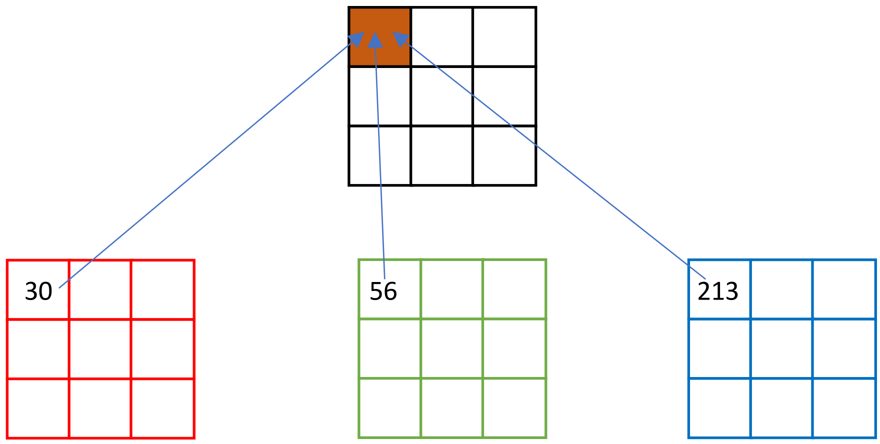
Each RGB image provided in the training set is a matrix of size 24 x 24 x 3. In other words, each image contains:

- One matrix for **RED** intensities of size 24 x 24 x 1
- One matrix for **GREEN** intensities of size 24 x 24 x 1
- One matrix for **BLUE** intensities of size 24 x 24 x 1

In RGB images each pixel is created by the combination of different amounts of red, green and blue. These amounts (intensities) can take values from 0 to 255.

Example:

Consider RGB images of 3 x 3 pixels. The final color that we observe in a pixel is a combination of red, green and blue



Possible features you can use:

- Smallest amount of **red** in the image (R_min)
- Mean of **red** values in the whole image (R_mean)
- Smallest amount of **green** in the image (G_min)
- Mean of **green** values in the whole image (G_mean)
- Smallest amount of **blue** in the image (B_min)
- Mean of **blue** values in the whole image (B_mean)

Step 2 – Calculate parameters $\theta, \mu_0, \mu_1, \Sigma$

See these calculations on page 9 of the slides.
Check also the Excel file where I explain how you make these calculations.

Step 3 – Test your Gaussian classifier using the Bayes rule

See the Bayes rule on page 5 of the slides.

Once you have calculated parameters $\theta, \mu_0, \mu_1, \Sigma$ you calculate the probability of an image x to belong to class 0, which is $p(y = 0|x)$:

$$p(y = 0|x) = \frac{p(x|y = 0)p(y = 0)}{p(x|y = 0)p(y = 0) + p(x|y = 1)p(y = 1))}$$

where

T means transpose

$$p(x|y = 0) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{\left(-\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)}$$

and

$|\Sigma|$ is the determinant of Σ

n is the number of features

$$p(y = 0) = 1 - \theta = \frac{30}{60} = 0.5$$

Then you calculate the probability that the image x belongs to class 1, which is $p(y = 1|x)$:

$$p(y = 1|x) = \frac{p(x|y = 1)p(y = 1)}{p(x|y = 0)p(y = 0) + p(x|y = 1)p(y = 1))}$$

where

$$p(x|y = 1) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{\left(-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)}$$

and

$$p(y = 1) = \theta = \frac{30}{60} = 0.5$$

The final decision/classification depends on the largest probability:

If $p(y = 1|x) > p(y = 0|x)$

the image x represents an image with a parasite (class 1)

Otherwise

it is an image without a parasite (class 0)