Image Data Augmentation

Deep Learning networks excel on the basis of the amount of training data that is fed into them. Thus, a powerful image classifier network can be built only with a huge training database of images.

But the obstacle that arises here is the limited number of images available, resulting in an imbalance of classes. The imbalance of classes leads to the neural network learning more features of one class whereas the leaving out the other, resulting in the lower data class less likely to get as the predicted class.

Image augmentation is primarily used in image-based Machine Learning and Deep Learning techniques due to these unique challenges that using images poses.

Image augmentation is used to create artificial images from a given set of images through using different ways of processing or multiple processing methods of these images such as flipping, rotation, colour changes, etc. Augmentation techniques applied are done in a way that preserves the features that are key in making the predictions, but rearranges the pixels in the images essentially creating a new image.

All these kinds of image properties change the way in which the image is perceived but not the contents or features of the image. Augmentation on the original images is used to obtain many variants of the same image. Thus, augmentation is a very useful technique to increase the training set size without capturing new images.

It is also used to force (overfit) the model to the available images which are increased by the augmentation process.

This section would explain the basics and the uses of each augmentation technique for the dataset creation.

Data Augmentations based on basic image manipulations

Geometric or Position transformations

Geometric or Position augmentation techniques are based on geometric transformations done through changing the pixel value positions in the original image to get a new image. Rotation, Flipping, Scaling, Cropping, Translation, Affine Transform are the majorly used techniques in Geometric augmentation and are explored.

Rotation

Image rotation is an operation that is used to rotate the image clockwise or counter-clockwise by a certain degree.

Rotation of an image is useful because of the seemingly different type of picture a rotation operation can produce. The content or features in the image are the same, but their positions in the rotation augmented image would be different. It thus gives rise to a whole set of new feature learnings for the convolutional neural networks. The more the amount of rotated images, the better the identification of images in any angle of capture.//

For a pixel at position (x_1,y_1) in an input image,

By rotating it with an angle θ about the origin of the input image, the new pixel position will be at (x_2,y_2)

$$x_2 = cos(\theta) * (x_1) + sin(\theta) * (y_1)$$

$$y_2 = -sin(\theta) * (x_1) + cos(\theta) * (y_1)$$

Rotations by smaller angles could be useful, but larger angle rotations may introduce more noise in the pictures rendering it useless for training purposes.

This augmentation technique is used to produce 9 images with different rotation angles. Thus, each original image produces 9 images, each with a different angle of rotation.

The angle of rotation for each of the 9 images is incremented by 10 degrees.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Images Produced through Rotation Variation
Good Quality	428	9	3852
Defect Quality	354	9	3186
		Total	**7038**

Flipping

Image flipping is an operation that is used to flip the image about a vertical or a horizontal axis, usually passing through the middle of the image in their respective axis.

Similar to the rotation operation, flipping produces a seemingly different image, but with the content of the image remaining the same, only difference is in the position of pixels which are changed.

For a pixel at position (x_1,y_1) in an input image,

(i) By reflecting it about a vertical axis of abscissa x_0 in the input image, the new pixel position will be at (x_2,y_2)

$$x_2 = -x_1 + (2 * x_0)$$

$$y_2 = y_1$$

(ii) By reflecting it about a horizontal axis of ordinate y_0 in the input image, the new pixel position will be at (x_2,y_2)

$$x_2 = x_1$$

$$y_2 = -y_1 + (2 * y_0)$$

This augmentation technique is used to produce 2 images with different flipping directions. Thus, each original image produces 2 images, each with a different type of flip. One in the horizontal and one in the vertical direction.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Images Produced through Flipping Variation
Good Quality	428	2	856
Defect Quality	354	2	708
		Total	**1564**

Cropping

Cropping is the removal of a certain number of pixels in an image.

Cropping results in images losing details due to the removal of pixels and thus reduces the size of the image. It can be used to remove unnecessary regions from the images and focus on regions that matter more for the convolutional neural network algorithm. Thus, unlike the other geometric transformations cropping results in loss of pixels and thus image detail.

This augmentation technique is used to produce 4 images with varying amounts of image cropping levels applied. This changes the level of details at each level and thus the features in each image.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Images Produced through Cropping Variation
Good Quality	428	4	1712
Defect Quality	354	4	1416
		Total	**3128**

Scaling

Scaling, usually means resizing an image by the same amount in both the horizontal and vertical directions.

The upscaling and downscaling of a lower resolution or a higher resolution image respectively, can be done by applying a suitable reconstruction filter 2D anti-aliasing filter respectively.

Scaling up an image can be very useful to “zoom” or highlight into features that the convolutional neural network should pick up as differentiating features.

This augmentation technique is used to produce 2 images with two levels of upscaling on the original image applied.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Images Produced through Scaling Variation
Good Quality	428	2	856
Defect Quality	354	2	708
		Total	**1564**

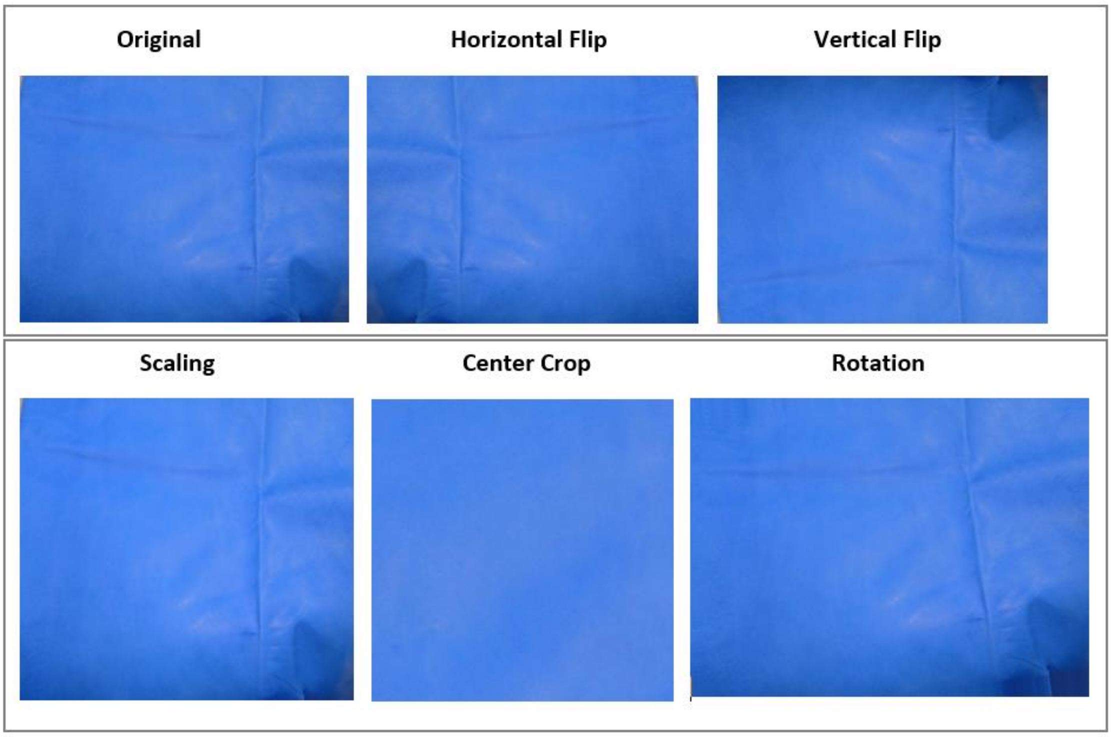

Fig 1. Examples of Geometric Augmentations

Translation

In image translation, the entire image is moved in the x,y direction along some direction.

In the translation operation, a pixel located at x_1,y_1 position in the original image is shifted to a new position x_2,y_2 in the augmented image. This is displaced by a user specified translation β_x and β_y.

$$x_2 = x_1 + \beta_x$$

$$y_2 = y_1 + \beta_y$$

The locations opposite to the image's direction of motion are filled with constant or random pixel values. Usually it is black pixels which are filled. Due to this introduction of undesirable noise, which will lead to erroneous detections, it is not used in the augmentation process.

Translation is generally discarded in favour of better feature highlights possible in cropping and scaling, since translation's main use case is moving an image's object to a different location for better focus, but that function is already given in cropping and scaling. Also, in these images, translation creates random pixel values fill to compensate for the moved pixels, which would decrease the accuracy of prediction due to wrong training data being produced.

Affine Transform

Affine Transform is an 2D geometric transformation which maps pixel values at x_1,y_1 position in the original image to x_2,y_2 a new position in the augmented image. This is done by applying a linear combination of translation, rotation, scaling and or shearing operations.

Affine Transform's Equation

$$\begin{vmatrix} x2 \\ y2 \end{vmatrix} = A \times \begin{vmatrix} x1 \\ y1 \end{vmatrix} + B$$

Where,

A and B are matrices that determine translation, scaling and rotation collectively.

Since, affine transform's main use case is for obtaining spatial or geometrically accurate information of the images, it is not essential in this case where the images have been taken almost with maximizing their use as it is, i.e, no further twisting or turning is necessary to get the required information.

Colour Augmentation or colour space transformations

Images are generally represented in the Red, Blue, Green colour space model. By this model, the images are encoded as size * width * 3 matrices of pixel values. Here, 3 is used to represent the three colour channels R,G,B.

Colour Augmentation is the process of manipulating these colour space pixel values so as to obtain a new coloured image without modifying the features.

The main use of using colour augmentation is for overcome lightning biases that challenge image recognition. Thus, colour distribution of the images can be altered for increasing the test data on various different lighting and coloured situations.

Grayscale

Grayscale images are images in which there is only one channel of colour, gray monochrome in them.
The pixels in the image denote the intensity of light.
In general Grayscale images are converted from RGB images by the following formula,

Grayscale Pixel of the corresponding pixel = (Red Component + Green Component Blue Component) of the corresponding pixel / 3

In this way, all the pixels are computed for the entire image.
It is the average of the corresponding R,G,B pixel values

In this augmentation process, the original RGB images are converted to grayscale images.
Thus each original image produces a grayscale image through grayscale augmentation.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Grayscale Images Produced
Good Quality	428	1	428
Defect Quality	354	1	354
		Total	**782**

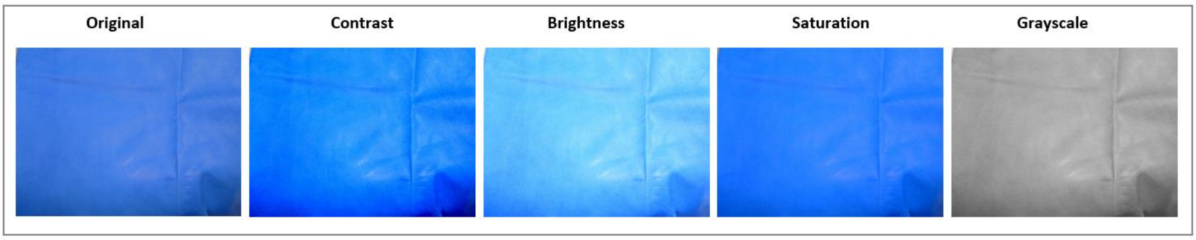

Fig 2. Examples of Colour Augmentations

Brightness

Brightness, a subjective measure can be generally defined as the overall lightness or darkness of the image.
It is a parameter that depends on the pixel value intensities of the 3 channels, R,G,B. If it’s more, it is considered brighter as it approaches more towards the white intensity of the 3 channels.

Basically, brightness adjustment is adding or subtracting equal pixel intensity values throughout the 3 channels.

Pixel Value at a position in each Channel = Pixel Value at the position in that Channel + Some Fixed Specified Value

In this way, the pixel values are calculated in all the channels throughout the image.

This augmentation technique is used to produce 2 images with different brightness levels.
Thus each original image produces two images, each with a different level of brightness.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Images Produced through Brightness Variation
Good Quality	428	2	856
Defect Quality	354	2	708
		Total	**1564**

Contrast

Contrast is the difference in the brightness levels between objects, background levels in an image.
In a high contrast image, details are usually highlighted.

This augmentation technique is used to produce 2 images with different contrast levels. Thus each original image produces two images, each with a different level of contrast.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Images Produced through Contrast Variation
Good Quality	428	2	856
Defect Quality	354	2	708
		Total	**1564**

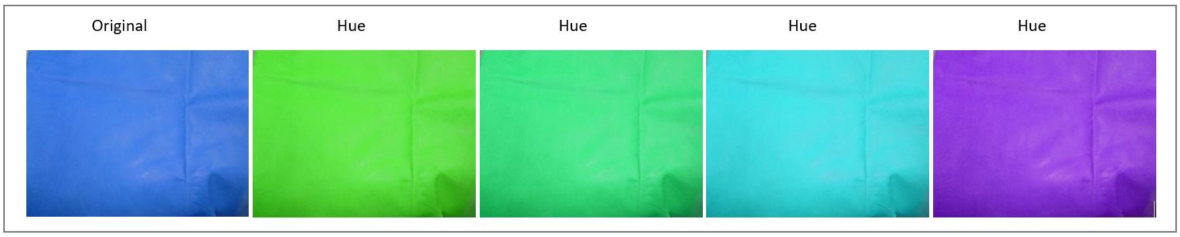

Fig 3. Examples of Colour Augmentations (Hue)

Saturation

Saturation usually describes the intensity of colour in an image.
This augmentation technique is used to produce 2 images with different saturation levels. Thus each original image produces two images, each with a different level of saturation.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Images Produced through Saturation Variation
Good Quality	428	2	856
Defect Quality	354	2	708

Total **1564**

Hue

Hue is often expressed as the dominant colour in the image.
This augmentation technique is used to produce 13 images with different hues. Thus each original image produces 13 images, each with a different level of hue.

Category of Leather Image	Number Of Images	Augmentation Multiplier	No of Images Produced through Hue Variation
Good Quality	428	13	5564
Defect Quality	354	13	4602

Total **10166**