Object Detection Using MASK-RCNN With Transfer Learning

1. Dataset

Our data set consists of 4 different types of identity cards. These include

- Pakistan passport 12 Images
- Pakistan CNIC 25 Images
- Singapore Passport 11 Images
- United Kingdom Passport 11 Images

All images are color images and consist of 3 channels RGB: red, green and blue. All of this data was collected from Google Images search results hence, sourced from various places all over the internet. Due to identity theft and other privacy and security concerns, the information of identity cards is sensitive information and not much data exists on it.

Dataset after Augmentation

Since, the dataset we have is very limited, we need to generate more images through data augmentation. The techniques for it will be discussed in later sections.

Using data augmentation, we generate 100 images from a single image and obtain the following results.

- Pakistan passport 1200 Images
- Pakistan CNIC 2500 Images
- Singapore Passport 1100 Images
- United Kingdom Passport 1100 Images

Training and Test Split

We split the data into training and test data with a 70-30 ratio obtaining the following split

	Total Data	Training Data	Test Data
Pakistan passport	1200	840	360
Pakistan CNIC	2500	1750	750
Singapore Passport	1100	770	330
UK Passport	1100	770	330

2. Annotations

We annotate each of the identity card with the following labels

Pakistan passport

- Surname
- Given Name
- Nationality
- Date of Birth
- Gender
- Father Name

Pakistan CNIC

- Name
- Father Name
- Gender
- Date of Birth
- Identity Number
- Country of Status

Singapore Passport

- Name
- Nationality
- Date of Birth
- Gender
- Place of Birth
- National ID

UK Passport

- Surname
- First Name
- Nationality
- Date of Birth
- Gender
- Place of Birth

Annotation Technique

For annotation, we made use of makesense.ai which is an open source and free to use image annotator which intelligently annotates the images using AI.

All of the images before augmentation were annotated manually on this online software. For annotating, a rectangle was drawn around each of the region of interest and it was labeled with a name. 6 labels were introduced in each image as discussed above. The annotations of each image were saved in the VOC XML format.

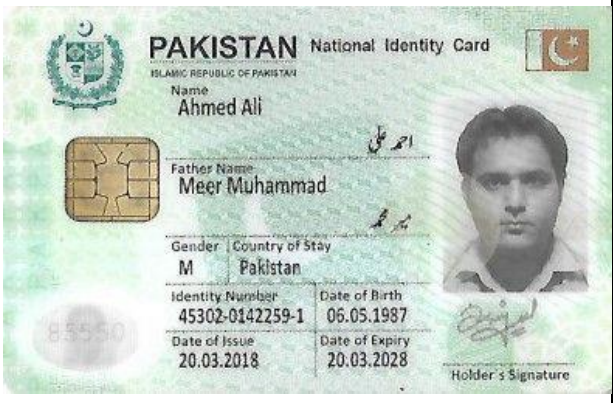
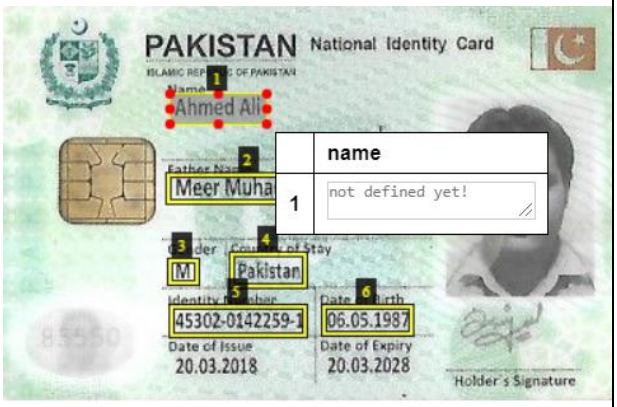
Annotation Files

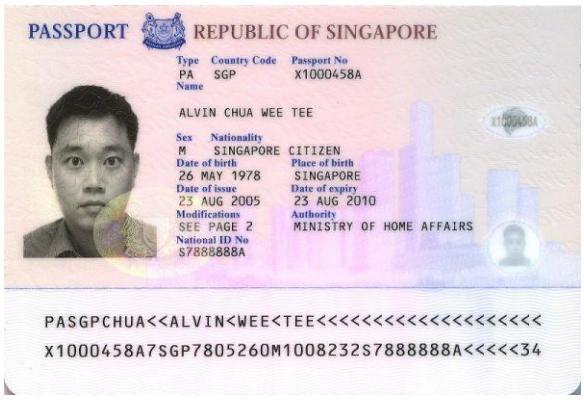
Each XML file contains complete information about that image. This include, the folder in which the images are placed, filename, Path to image, width, height, depth, and object tags for all the labels. A sample Object tag is shown as follows

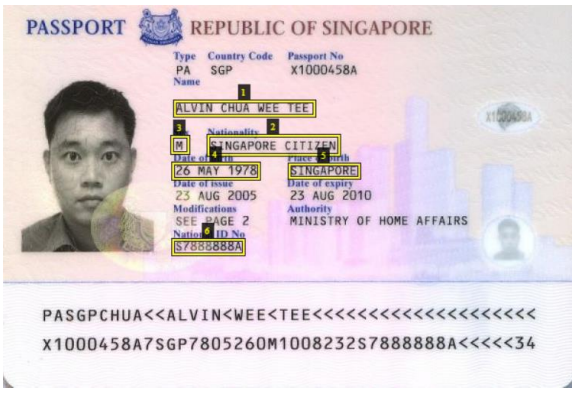
```
<object>
      <name>Surname</name>
      <pose>Unspecified</pose>
      <truncated>Unspecified</truncated>
      <difficult>Unspecified</difficult>
      <bndbox>
            <xmin>148</xmin>
            <ymin>56</ymin>
            <xmax>196</xmax>
            <ymax>71</ymax>
      </bndbox>
</object>
```

<bndbox> block indicated the bounding box of the image. xmin, ymin represent the top left corner and xmax, ymax indicate the bottom right corner coordinates of the bounding box in the picture. <name> block is the label of the bounding box.

Some Original and Annotated Images

Original Image	Annotated Image
	





3. Augmentation Methods

For the purpose of augmentation, we use the **imgaug** library, imgaug allows us to tune many parameters to augment the images in various ways.

We selected 7 of these parameters for augmentation

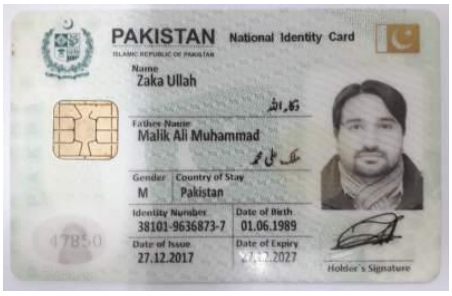
1. Scope
2. Rotate
3. Translate the image in x/y direction
4. Flip
5. Zoom
6. Gaussian Blur
7. Brightness

We have adjusted certain parameters for these functions. For each new image to be augmented, the function choses any 2 of these 7 functions. The selected two filters are applied on the image and the result is a new augmented image generated.

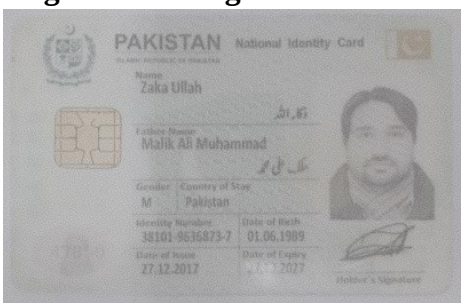
Before augmentation, we first resize the image. All images are resized such that the pictures' width and height are 600px or less. To resize the images a well, we use imgaug library.

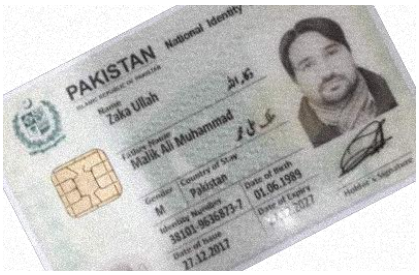
Some Original and Augmented Images

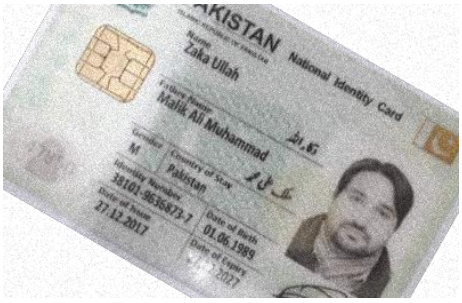
Original Image

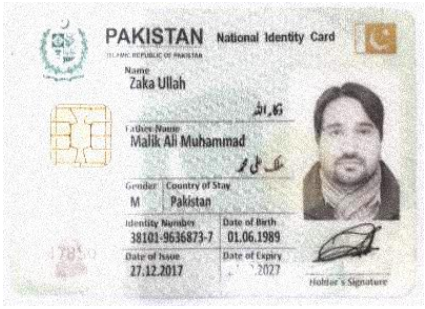


Augmented Image









In the above images, it can be clearly observed one image has been rotated and zoomed. One of them rotated and Gaussian blur introduced. One is brightened and the brightne3ss is reduced in the 4th one.

4. Augmented Data Annotation

Annotation Technique

The augmented data needs to be annotated as well so that the model can be trained on it. This done in such a way that when the image is augmented, corresponding object bounding boxes are also augmented accordingly.

The code written for this purpose, clip the bounding box if the object of interest is partially outside of image pane and ignores image altogether if no bounding boxes left in image pane.

The code provided passes the resized image to be augmented and saved to the destined folder. The corresponding bounding boxes annotation for each of these augmentations is stored in a dataframe.

Annotation Files

This dataframe for image annotations is correspondingly converted to a csv file. This csv file contains all the parameters and bounding boxes of augmented images. Example of it is shown below.

A	B	C	D	E	F	G	H
aug92_0710-Datapage.jpg	1200	854	UKPassport	151.7982	52.0805	773.1105	479.9674
aug92_0820-Biodata.jpg	1200	818	Surname	724.2625	0	779.7625	15.04735
aug92_0820-Biodata.jpg	1200	818	FirstName	626.2625	26.04736	777.7625	46.04736
aug92_0820-Biodata.jpg	1200	818	Nationality	625.2625	58.04736	777.7625	77.04736
aug92_0820-Biodata.jpg	1200	818	DOB	765.2625	86.04736	781.2625	102.5474
aug92_0820-Biodata.jpg	1200	818	Gender	674.7625	86.04736	737.2625	106.0474
aug92_0820-Biodata.jpg	1200	818	PlaceOfBirth	389.2625	0	980.7625	207.0473
aug92_0820-Biodata.jpg	1200	818	UKPassport				
aug92_0830-Datapage.jpg	1200	823	Surname	933.262	81.05819	1022.152	104.4191
aug92_0830-Datapage.jpg	1200	823	FirstName	858.2133	109.483	1019.354	134.3542
aug92_0830-Datapage.jpg	1200	823	Nationality	834.3644	138.2568	1018.691	168.3988
aug92_0830-Datapage.jpg	1200	823	DOB	847.9849	166.5777	1020.663	193.0845
aug92_0830-Datapage.jpg	1200	823	Gender	997.5365	205.4465	1016.478	223.9365
aug92_0830-Datapage.jpg	1200	823	PlaceOfBirth	769.7469	192.8372	958.0674	223.2004
aug92_0830-Datapage.jpg	1200	823	UKPassport	601.7166	2.129298	1200	413.3874
aug92_1310-Biodata.jpg	1200	814	Surname	235.4295	76.83117	312.6207	98.04247
aug92_1310-Biodata.jpg	1200	814	FirstName	237.652	106.7546	291.322	125.9754
aug92_1310-Biodata.jpg	1200	814	Nationality	237.3768	130.2022	423.4288	155.0128
aug92_1310-Biodata.jpg	1200	814	DOB	235.04	158.184	412.684	184.6342
aug92_1310-Biodata.jpg	1200	814	Gender	241.4278	197.8835	254.5284	212.8979
aug92_1310-Biodata.jpg	1200	814	PlaceOfBirth	298.3348	190.1344	414.5335	213.9901
aug92_1310-Biodata.jpg	1200	814	UKPassport	40.29538	0	650.1285	397.404
aug92_1528277944.perfect_20passport.jpg	500	338	Surname	167.6166	53.34026	327.9602	74.70633
aug92_1528277944.perfect_20passport.jpg	500	338	FirstName	167.2066	82.63995	266.4584	97.72229
aug92_1528277944.perfect_20passport.jpg	500	338	Nationality	165.6704	104.0242	296.0417	125.7582
aug92_1528277944.perfect_20passport.jpg	500	338	DOB	169.2805	133.1863	288.5069	147.6901