Adaptive Neural Architectures for Cross-Domain Object Detection in Changing Environments

Data Collection

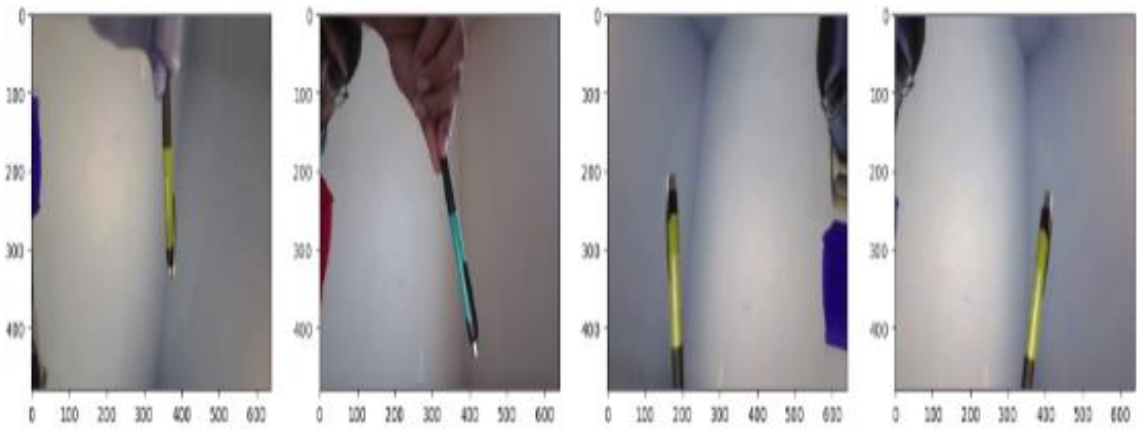


Fig: Sample from Data Collection

Data Augmentation

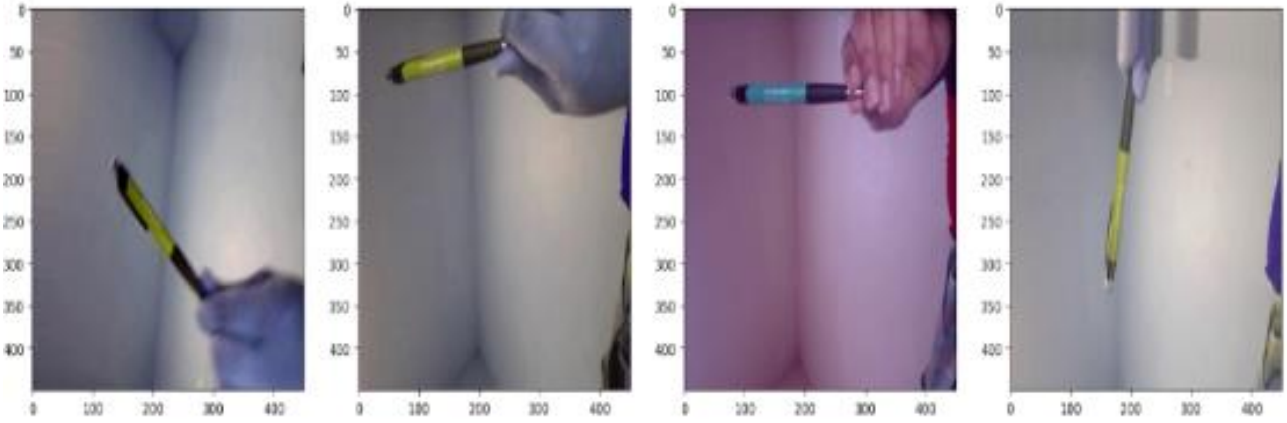


Fig: Result of Data Augmentation

Model Building

Model: "vgg16"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, None, None, 3)	0
block1_conv1 (Conv2D)	(None, None, None, 64)	1,792
block1_conv2 (Conv2D)	(None, None, None, 64)	36,928
block1_pool (MaxPooling2D)	(None, None, None, 64)	0
block2_conv1 (Conv2D)	(None, None, None, 128)	73,856
block2_conv2 (Conv2D)	(None, None, None, 128)	147,584
block2_pool (MaxPooling2D)	(None, None, None, 128)	0
block3_conv1 (Conv2D)	(None, None, None, 256)	295,168
block3_conv2 (Conv2D)	(None, None, None, 256)	590,080
block3_conv3 (Conv2D)	(None, None, None, 256)	590,080
block3_pool (MaxPooling2D)	(None, None, None, 256)	0
block4_conv1 (Conv2D)	(None, None, None, 512)	1,180,160
block4_conv2 (Conv2D)	(None, None, None, 512)	2,359,808
block4_conv3 (Conv2D)	(None, None, None, 512)	2,359,808
block4_pool (MaxPooling2D)	(None, None, None, 512)	0
block5_conv1 (Conv2D)	(None, None, None, 512)	2,359,808
block5_conv2 (Conv2D)	(None, None, None, 512)	2,359,808
block5_conv3 (Conv2D)	(None, None, None, 512)	2,359,808
block5_pool (MaxPooling2D)	(None, None, None, 512)	0

Total params: 14,714,688 (56.13 MB)
Trainable params: 14,714,688 (56.13 MB)
Non-trainable params: 0 (0.00 B)

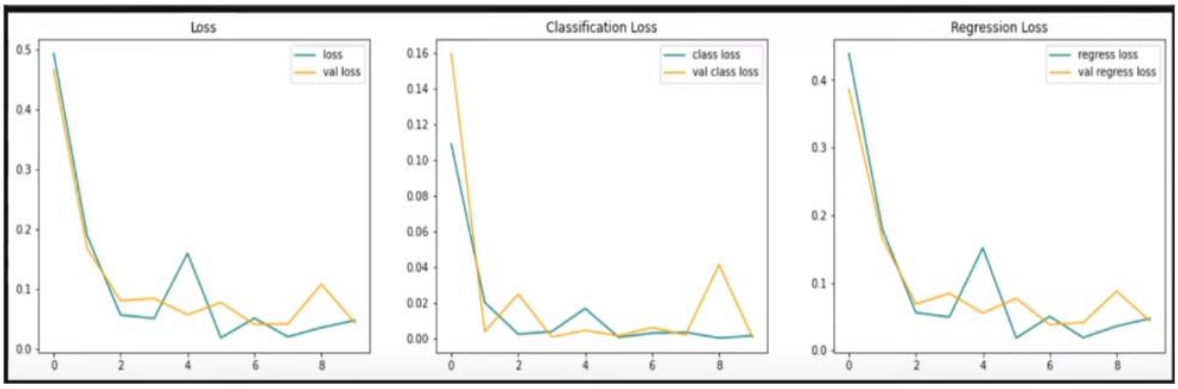
Model: "functional_1"

Layer (type)	Output Shape	Param #	Connected to
input_layer_1 (InputLayer)	(None, 120, 120, 3)	0	-
vgg16 (Functional)	(None, 3, 3, 512)	14,714,688	input_layer_1[0]…
global_max_pooling… (GlobalMaxPooling2…	(None, 512)	0	vgg16[0][0]
global_max_pooling… (GlobalMaxPooling2…	(None, 512)	0	vgg16[0][0]
dense (Dense)	(None, 2048)	1,050,624	global_max_pooli…
dense_2 (Dense)	(None, 2048)	1,050,624	global_max_pooli…
dense_1 (Dense)	(None, 1)	2,049	dense[0][0]
dense_3 (Dense)	(None, 4)	8,196	dense_2[0][0]

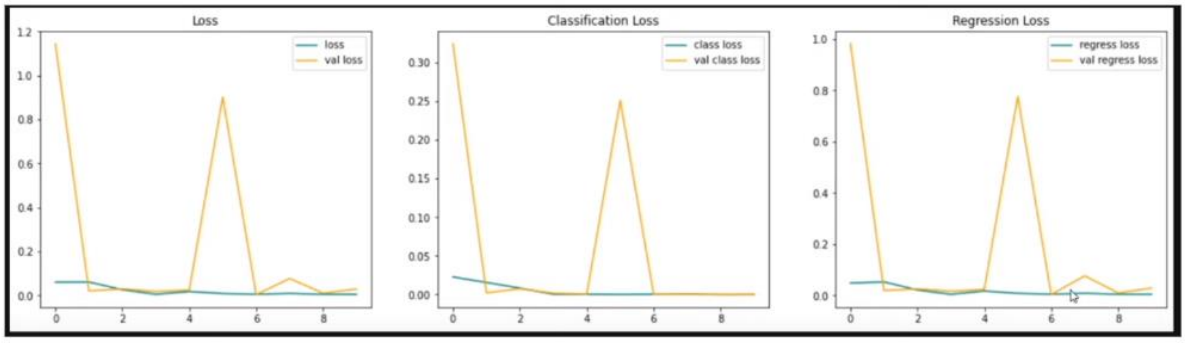
Total params: 16,826,181 (64.19 MB)
Trainable params: 16,826,181 (64.19 MB)
Non-trainable params: 0 (0.00 B)

Fig: Instance of our Network

Model Training



Model Evaluation



Output: