DSC Workshop: Object Detection

Joshua S Raju

Dependencies

In [60]:
```
!nvidia-smi
```

```
Sat Feb 22 17:13:46 2025
+-----------------------------------------------------------------------------------------+
| NVIDIA-SMI 560.35.03              Driver Version: 560.35.03      CUDA Version: 12.6     |
|-----------------------------------------+------------------------+----------------------+
| GPU  Name                 Persistence-M | Bus-Id          Disp.A | Volatile Uncorr. ECC |
| Fan  Temp   Perf          Pwr:Usage/Cap |           Memory-Usage | GPU-Util  Compute M. |
|                                         |                        |               MIG M. |
|=========================================+========================+======================|
|   0  Tesla P100-PCIE-16GB           Off |   00000000:00:04.0 Off |                    0 |
| N/A   35C    P0             31W /  250W |     619MiB /  16384MiB |      0%      Default |
|                                         |                        |                  N/A |
+-----------------------------------------+------------------------+----------------------+

+-----------------------------------------------------------------------------------------+
| Processes:                                                                              |
|  GPU   GI   CI        PID   Type   Process name                              GPU Memory |
|        ID   ID                                                               Usage      |
|=========================================================================================|
+-----------------------------------------------------------------------------------------+
```

In [61]:
```
!pip install ultralytics roboflow
```

```
Requirement already satisfied: ultralytics in /usr/local/lib/python3.10/dist-pack
ages (8.3.78)
Requirement already satisfied: roboflow in /usr/local/lib/python3.10/dist-package
s (1.1.54)
Requirement already satisfied: numpy<=2.1.1,>=1.23.0 in /usr/local/lib/python3.1
0/dist-packages (from ultralytics) (1.26.4)
Requirement already satisfied: matplotlib>=3.3.0 in /usr/local/lib/python3.10/dis
t-packages (from ultralytics) (3.7.5)
Requirement already satisfied: opencv-python>=4.6.0 in /usr/local/lib/python3.10/
dist-packages (from ultralytics) (4.10.0.84)
Requirement already satisfied: pillow>=7.1.2 in /usr/local/lib/python3.10/dist-pa
ckages (from ultralytics) (11.0.0)
Requirement already satisfied: pyyaml>=5.3.1 in /usr/local/lib/python3.10/dist-pa
ckages (from ultralytics) (6.0.2)
Requirement already satisfied: requests>=2.23.0 in /usr/local/lib/python3.10/dist
-packages (from ultralytics) (2.32.3)
Requirement already satisfied: scipy>=1.4.1 in /usr/local/lib/python3.10/dist-pac
kages (from ultralytics) (1.13.1)
Requirement already satisfied: torch>=1.8.0 in /usr/local/lib/python3.10/dist-pac
kages (from ultralytics) (2.5.1+cu121)
Requirement already satisfied: torchvision>=0.9.0 in /usr/local/lib/python3.10/di
st-packages (from ultralytics) (0.20.1+cu121)
Requirement already satisfied: tqdm>=4.64.0 in /usr/local/lib/python3.10/dist-pac
kages (from ultralytics) (4.67.1)
Requirement already satisfied: psutil in /usr/local/lib/python3.10/dist-packages
(from ultralytics) (5.9.5)
Requirement already satisfied: py-cpuinfo in /usr/local/lib/python3.10/dist-packa
ges (from ultralytics) (9.0.0)
Requirement already satisfied: pandas>=1.1.4 in /usr/local/lib/python3.10/dist-pa
ckages (from ultralytics) (2.2.3)
Requirement already satisfied: seaborn>=0.11.0 in /usr/local/lib/python3.10/dist-
packages (from ultralytics) (0.12.2)
Requirement already satisfied: ultralytics-thop>=2.0.0 in /usr/local/lib/python3.
10/dist-packages (from ultralytics) (2.0.14)
Requirement already satisfied: certifi in /usr/local/lib/python3.10/dist-packages
(from roboflow) (2025.1.31)
Requirement already satisfied: idna==3.7 in /usr/local/lib/python3.10/dist-packag
es (from roboflow) (3.7)
Requirement already satisfied: cycler in /usr/local/lib/python3.10/dist-packages
(from roboflow) (0.12.1)
Requirement already satisfied: kiwisolver>=1.3.1 in /usr/local/lib/python3.10/dis
t-packages (from roboflow) (1.4.7)
Requirement already satisfied: opencv-python-headless==4.10.0.84 in /usr/local/li
b/python3.10/dist-packages (from roboflow) (4.10.0.84)
Requirement already satisfied: python-dateutil in /usr/local/lib/python3.10/dist-
packages (from roboflow) (2.9.0.post0)
Requirement already satisfied: python-dotenv in /usr/local/lib/python3.10/dist-pa
ckages (from roboflow) (1.0.1)
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (fr
om roboflow) (1.17.0)
Requirement already satisfied: urllib3>=1.26.6 in /usr/local/lib/python3.10/dist-
packages (from roboflow) (2.3.0)
Requirement already satisfied: requests-toolbelt in /usr/local/lib/python3.10/dis
t-packages (from roboflow) (1.0.0)
Requirement already satisfied: filetype in /usr/local/lib/python3.10/dist-package
s (from roboflow) (1.2.0)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist
-packages (from matplotlib>=3.3.0->ultralytics) (1.3.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dis
t-packages (from matplotlib>=3.3.0->ultralytics) (4.55.3)
```

```
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-
packages (from matplotlib>=3.3.0->ultralytics) (24.2)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist
-packages (from matplotlib>=3.3.0->ultralytics) (3.2.0)
Requirement already satisfied: mkl_fft in /usr/local/lib/python3.10/dist-packages
(from numpy<=2.1.1,>=1.23.0->ultralytics) (1.3.8)
Requirement already satisfied: mkl_random in /usr/local/lib/python3.10/dist-packa
ges (from numpy<=2.1.1,>=1.23.0->ultralytics) (1.2.4)
Requirement already satisfied: mkl_umath in /usr/local/lib/python3.10/dist-packag
es (from numpy<=2.1.1,>=1.23.0->ultralytics) (0.1.1)
Requirement already satisfied: mkl in /usr/local/lib/python3.10/dist-packages (fr
om numpy<=2.1.1,>=1.23.0->ultralytics) (2025.0.1)
Requirement already satisfied: tbb4py in /usr/local/lib/python3.10/dist-packages
(from numpy<=2.1.1,>=1.23.0->ultralytics) (2022.0.0)
Requirement already satisfied: mkl-service in /usr/local/lib/python3.10/dist-pack
ages (from numpy<=2.1.1,>=1.23.0->ultralytics) (2.4.1)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-pac
kages (from pandas>=1.1.4->ultralytics) (2025.1)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.10/dist-p
ackages (from pandas>=1.1.4->ultralytics) (2025.1)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python
3.10/dist-packages (from requests>=2.23.0->ultralytics) (3.4.1)
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-package
s (from torch>=1.8.0->ultralytics) (3.17.0)
Requirement already satisfied: typing-extensions>=4.8.0 in /usr/local/lib/python
3.10/dist-packages (from torch>=1.8.0->ultralytics) (4.12.2)
Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-package
s (from torch>=1.8.0->ultralytics) (3.4.2)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages
(from torch>=1.8.0->ultralytics) (3.1.4)
Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-packages
(from torch>=1.8.0->ultralytics) (2024.9.0)
Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.10/dist-pa
ckages (from torch>=1.8.0->ultralytics) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.10/di
st-packages (from sympy==1.13.1->torch>=1.8.0->ultralytics) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-
packages (from jinja2->torch>=1.8.0->ultralytics) (3.0.2)
Requirement already satisfied: intel-openmp>=2024 in /usr/local/lib/python3.10/di
st-packages (from mkl->numpy<=2.1.1,>=1.23.0->ultralytics) (2024.2.0)
Requirement already satisfied: tbb==2022.* in /usr/local/lib/python3.10/dist-pack
ages (from mkl->numpy<=2.1.1,>=1.23.0->ultralytics) (2022.0.0)
Requirement already satisfied: tcmlib==1.* in /usr/local/lib/python3.10/dist-pack
ages (from tbb==2022.*->mkl->numpy<=2.1.1,>=1.23.0->ultralytics) (1.2.0)
Requirement already satisfied: intel-cmplr-lib-rt in /usr/local/lib/python3.10/di
st-packages (from mkl_umath->numpy<=2.1.1,>=1.23.0->ultralytics) (2024.2.0)
Requirement already satisfied: intel-cmplr-lib-ur==2024.2.0 in /usr/local/lib/pyt
hon3.10/dist-packages (from intel-openmp>=2024->mkl->numpy<=2.1.1,>=1.23.0->ultra
lytics) (2024.2.0)
```

In [62]:

```
import cv2
import matplotlib.pylab as plt

import os

from sklearn.metrics.pairwise import cosine_similarity

from ultralytics import YOLO
```

```
import tensorflow as tf
import torch
```

Understanding the Model

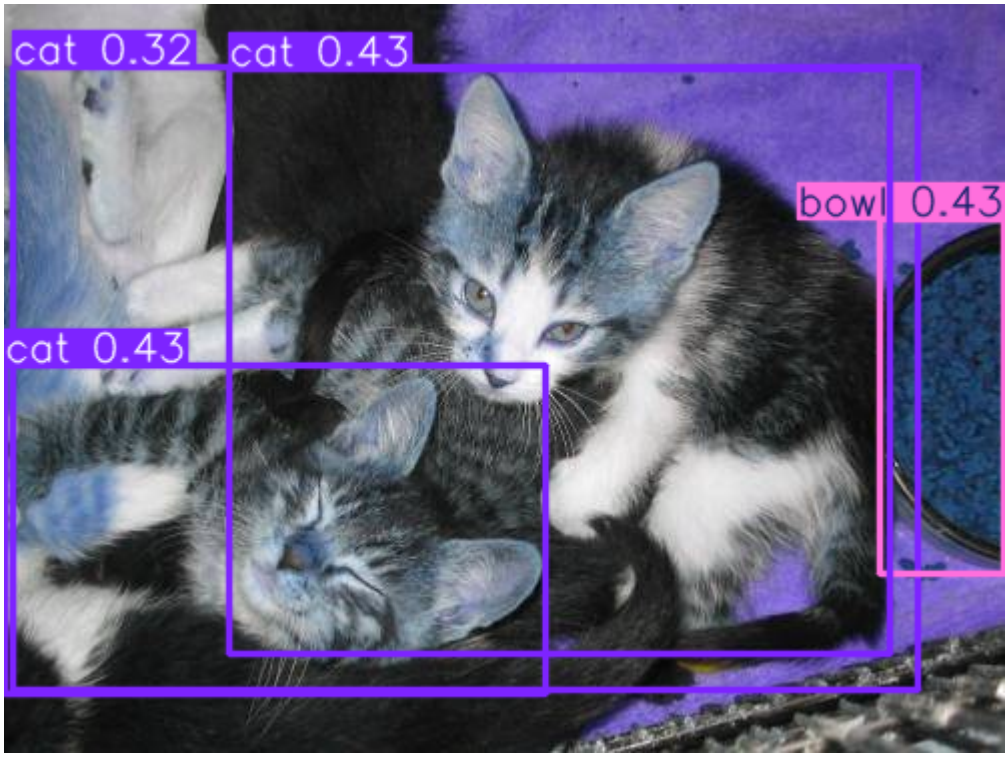
In [63]:
```
# Loading the model
model = YOLO('yolo11n.pt')
```

In [64]:
```
# Test Img
img_path = '/kaggle/input/cat-and-dog/training_set/training_set/cats/cat.1025.jp
img = cv2.cvtColor(cv2.imread(img_path), cv2.COLOR_BGR2RGB)
plt.imshow(img)
```

Out[64]:
```
<matplotlib.image.AxesImage at 0x79009419faf0>
```

In [65]:
```
# Predicting Img
results = model.predict(img)
results[0].show()
```

```
0: 480x640 3 cats, 1 bowl, 12.5ms
Speed: 2.4ms preprocess, 12.5ms inference, 1.7ms postprocess per image at shape
(1, 3, 480, 640)
```



In [66]:
```
model.names
```

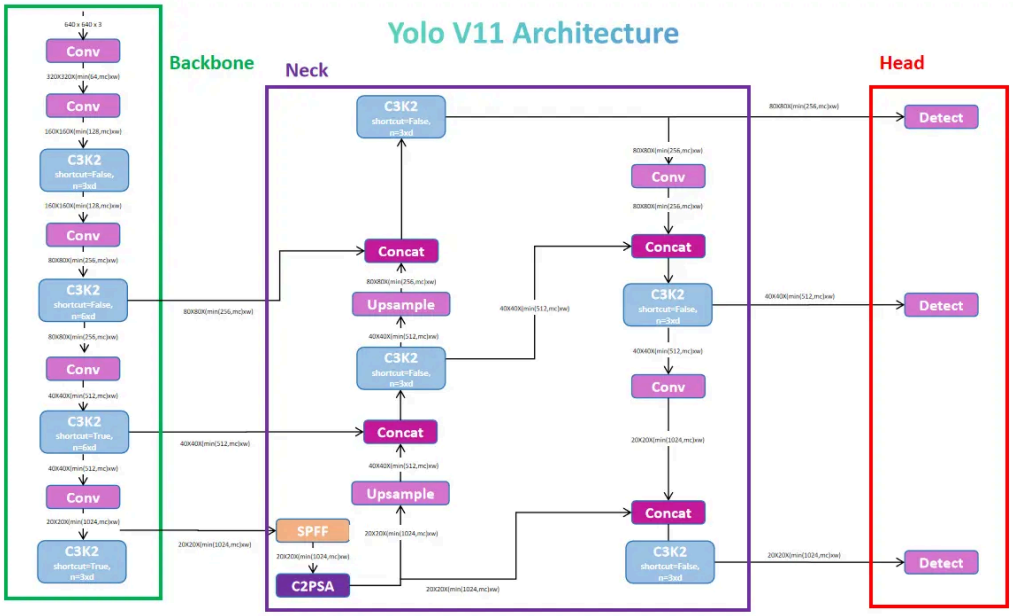
Out[66]:
```
{0: 'person',
 1: 'bicycle',
 2: 'car',
 3: 'motorcycle',
 4: 'airplane',
 5: 'bus',
 6: 'train',
 7: 'truck',
 8: 'boat',
 9: 'traffic light',
 10: 'fire hydrant',
 11: 'stop sign',
 12: 'parking meter',
 13: 'bench',
 14: 'bird',
 15: 'cat',
 16: 'dog',
 17: 'horse',
 18: 'sheep',
 19: 'cow',
 20: 'elephant',
 21: 'bear',
 22: 'zebra',
 23: 'giraffe',
 24: 'backpack',
 25: 'umbrella',
 26: 'handbag',
 27: 'tie',
 28: 'suitcase',
 29: 'frisbee',
 30: 'skis',
 31: 'snowboard',
 32: 'sports ball',
 33: 'kite',
 34: 'baseball bat',
 35: 'baseball glove',
 36: 'skateboard',
 37: 'surfboard',
 38: 'tennis racket',
 39: 'bottle',
 40: 'wine glass',
 41: 'cup',
 42: 'fork',
 43: 'knife',
 44: 'spoon',
 45: 'bowl',
 46: 'banana',
 47: 'apple',
 48: 'sandwich',
 49: 'orange',
 50: 'broccoli',
 51: 'carrot',
 52: 'hot dog',
 53: 'pizza',
 54: 'donut',
 55: 'cake',
 56: 'chair',
 57: 'couch',
 58: 'potted plant',
 59: 'bed',
```

```
 60: 'dining table',
 61: 'toilet',
 62: 'tv',
 63: 'laptop',
 64: 'mouse',
 65: 'remote',
 66: 'keyboard',
 67: 'cell phone',
 68: 'microwave',
 69: 'oven',
 70: 'toaster',
 71: 'sink',
 72: 'refrigerator',
 73: 'book',
 74: 'clock',
 75: 'vase',
 76: 'scissors',
 77: 'teddy bear',
 78: 'hair drier',
 79: 'toothbrush'}
```

In [67]:
```python
# Understanding the output
for result in results:
    boxes = result.boxes.xyxy
    conf = result.boxes.conf
    cls = result.boxes.cls
    for i in range(len(boxes)):
        print(cls[i])
        if conf[i] > 0.434:
            print(boxes[i])
            x1, y1, x2, y2 = map(int, boxes[i])
            img_crop = img[y1:y2, x1:x2]
            plt.imshow(img_crop)
```

```
tensor(15., device='cuda:0')
tensor([5.3549e-02, 1.8051e+02, 2.7061e+02, 3.4415e+02], device='cuda:0')
tensor(15., device='cuda:0')
tensor(45., device='cuda:0')
tensor(15., device='cuda:0')
```

Model Backbone



Low-Level Features

- Textures
- Edges
- Patterns
- Shapes
- More spatial

High-Level Features

- Semantic information - object tracking, recognition
- Less spatial

In [68]:
```
# Extracting the backbone
backbone = model.model.model[:10]
print(backbone)
```

```
Sequential(
  (0): Conv(
    (conv): Conv2d(3, 16, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
    (act): SiLU(inplace=True)
  )
  (1): Conv(
    (conv): Conv2d(16, 32, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
    (act): SiLU(inplace=True)
  )
  (2): C3k2(
    (cv1): Conv(
      (conv): Conv2d(32, 32, kernel_size=(1, 1), stride=(1, 1))
      (act): SiLU(inplace=True)
    )
    (cv2): Conv(
      (conv): Conv2d(48, 64, kernel_size=(1, 1), stride=(1, 1))
      (act): SiLU(inplace=True)
    )
    (m): ModuleList(
      (0): Bottleneck(
        (cv1): Conv(
          (conv): Conv2d(16, 8, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1))
          (act): SiLU(inplace=True)
        )
        (cv2): Conv(
          (conv): Conv2d(8, 16, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1))
          (act): SiLU(inplace=True)
        )
      )
    )
  )
  (3): Conv(
    (conv): Conv2d(64, 64, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
    (act): SiLU(inplace=True)
  )
  (4): C3k2(
    (cv1): Conv(
      (conv): Conv2d(64, 64, kernel_size=(1, 1), stride=(1, 1))
      (act): SiLU(inplace=True)
    )
    (cv2): Conv(
      (conv): Conv2d(96, 128, kernel_size=(1, 1), stride=(1, 1))
      (act): SiLU(inplace=True)
    )
    (m): ModuleList(
      (0): Bottleneck(
        (cv1): Conv(
          (conv): Conv2d(32, 16, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1))
          (act): SiLU(inplace=True)
        )
        (cv2): Conv(
          (conv): Conv2d(16, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1))
          (act): SiLU(inplace=True)
        )
      )
    )
```

```
    )
    (5): Conv(
      (conv): Conv2d(128, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
      (act): SiLU(inplace=True)
    )
    (6): C3k2(
      (cv1): Conv(
        (conv): Conv2d(128, 128, kernel_size=(1, 1), stride=(1, 1))
        (act): SiLU(inplace=True)
      )
      (cv2): Conv(
        (conv): Conv2d(192, 128, kernel_size=(1, 1), stride=(1, 1))
        (act): SiLU(inplace=True)
      )
      (m): ModuleList(
        (0): C3k(
          (cv1): Conv(
            (conv): Conv2d(64, 32, kernel_size=(1, 1), stride=(1, 1))
            (act): SiLU(inplace=True)
          )
          (cv2): Conv(
            (conv): Conv2d(64, 32, kernel_size=(1, 1), stride=(1, 1))
            (act): SiLU(inplace=True)
          )
          (cv3): Conv(
            (conv): Conv2d(64, 64, kernel_size=(1, 1), stride=(1, 1))
            (act): SiLU(inplace=True)
          )
          (m): Sequential(
            (0): Bottleneck(
              (cv1): Conv(
                (conv): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=
(1, 1))
                (act): SiLU(inplace=True)
              )
              (cv2): Conv(
                (conv): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=
(1, 1))
                (act): SiLU(inplace=True)
              )
            )
            (1): Bottleneck(
              (cv1): Conv(
                (conv): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=
(1, 1))
                (act): SiLU(inplace=True)
              )
              (cv2): Conv(
                (conv): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=
(1, 1))
                (act): SiLU(inplace=True)
              )
            )
          )
        )
      )
    )
    (7): Conv(
      (conv): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
      (act): SiLU(inplace=True)
```

```
    )
    (8): C3k2(
      (cv1): Conv(
        (conv): Conv2d(256, 256, kernel_size=(1, 1), stride=(1, 1))
        (act): SiLU(inplace=True)
      )
      (cv2): Conv(
        (conv): Conv2d(384, 256, kernel_size=(1, 1), stride=(1, 1))
        (act): SiLU(inplace=True)
      )
      (m): ModuleList(
        (0): C3k(
          (cv1): Conv(
            (conv): Conv2d(128, 64, kernel_size=(1, 1), stride=(1, 1))
            (act): SiLU(inplace=True)
          )
          (cv2): Conv(
            (conv): Conv2d(128, 64, kernel_size=(1, 1), stride=(1, 1))
            (act): SiLU(inplace=True)
          )
          (cv3): Conv(
            (conv): Conv2d(128, 128, kernel_size=(1, 1), stride=(1, 1))
            (act): SiLU(inplace=True)
          )
          (m): Sequential(
            (0): Bottleneck(
              (cv1): Conv(
                (conv): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=
(1, 1))
                (act): SiLU(inplace=True)
              )
              (cv2): Conv(
                (conv): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=
(1, 1))
                (act): SiLU(inplace=True)
              )
            )
            (1): Bottleneck(
              (cv1): Conv(
                (conv): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=
(1, 1))
                (act): SiLU(inplace=True)
              )
              (cv2): Conv(
                (conv): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=
(1, 1))
                (act): SiLU(inplace=True)
              )
            )
          )
        )
      )
    )
    (9): SPPF(
      (cv1): Conv(
        (conv): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1))
        (act): SiLU(inplace=True)
      )
      (cv2): Conv(
        (conv): Conv2d(512, 256, kernel_size=(1, 1), stride=(1, 1))
```

```
      (act): SiLU(inplace=True)
    )
    (m): MaxPool2d(kernel_size=5, stride=1, padding=2, dilation=1, ceil_mode=Fals
e)
  )
)
```

In [69]:
```
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
```

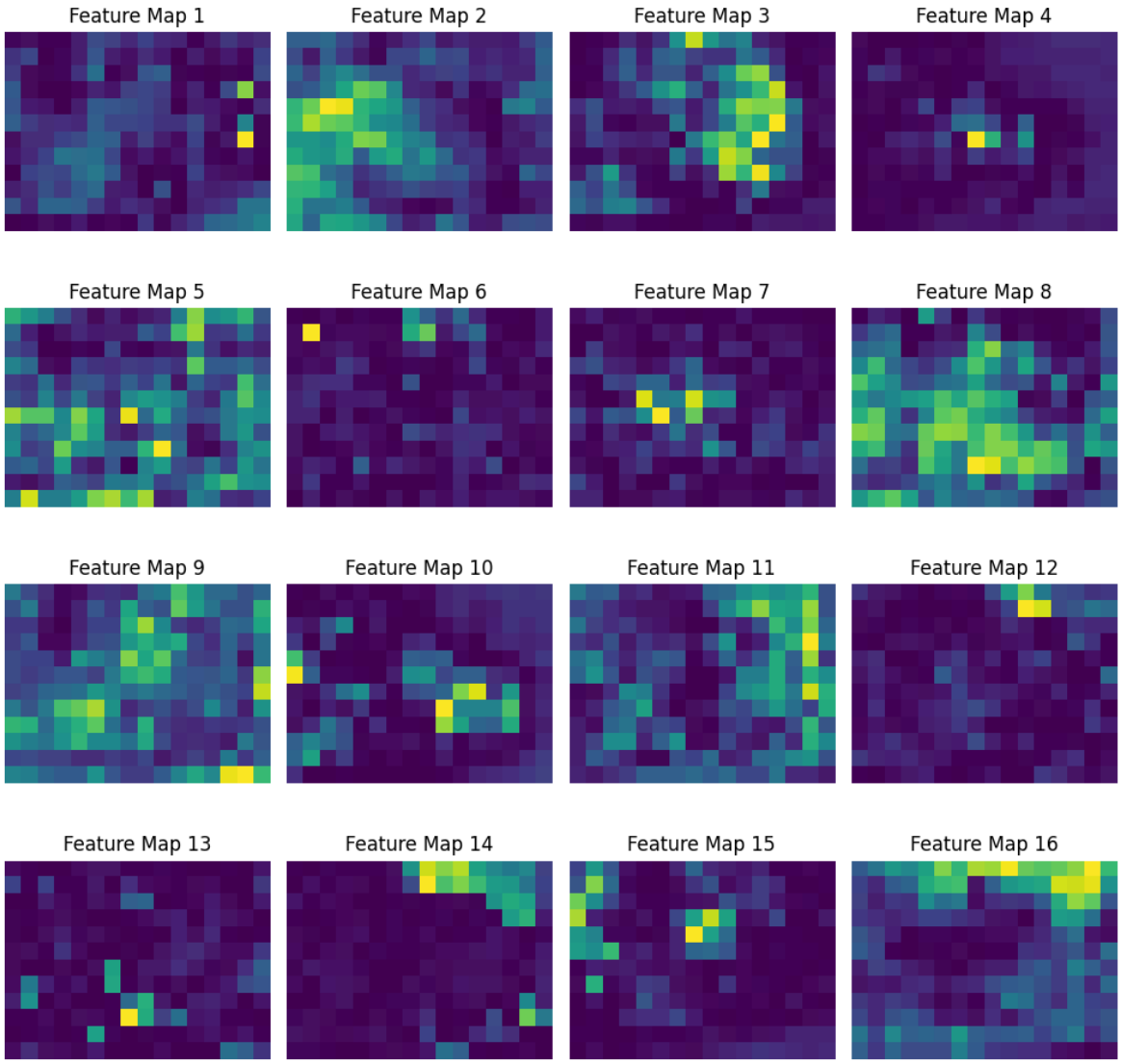
In [70]:
```
# Extarcting Test Img Feature Map
img_tensor = (torch.tensor(img).permute(2, 0, 1).float().unsqueeze(0)/255.0).to(
backbone(img_tensor)
```

Out[70]:
```
tensor([[[[-2.6978e-01, -2.5696e-01, -2.7331e-01,  ..., -1.9444e-01, -2.2651e-0
1, -2.1103e-01],
          [-1.9794e-01, -1.6513e-01, -2.7693e-01,  ..., -2.3948e-01, -2.6609e-0
1, -1.7218e-01],
          [-1.8171e-01, -2.6244e-01, -2.7185e-01,  ..., -2.7139e-01, -2.7400e-0
1, -1.5748e-01],
          ...,
          [-2.3416e-01, -2.1152e-01, -1.5165e-01,  ..., -2.2113e-01, -2.0727e-0
1, -1.6345e-01],
          [-1.4341e-01, -2.1936e-01, -1.0931e-01,  ..., -1.8965e-01, -8.2362e-0
2, -8.0038e-02],
          [-2.7823e-01, -2.7637e-01, -2.5666e-01,  ..., -9.6568e-02, -7.7649e-0
2, -4.1870e-02]],

         [[ 1.6667e-01,  1.8470e-01,  2.1821e-01,  ..., -1.5483e-01,  1.2709e-0
4,  1.4544e-01],
          [ 6.7708e-02, -1.6009e-01,  1.7270e-01,  ..., -1.2680e-01, -9.8899e-0
2,  2.0566e-01],
          [ 1.9675e-01,  1.0329e-02,  2.8604e-01,  ..., -1.2445e-01, -1.0793e-0
1,  1.6405e-01],
          ...,
          [ 6.7959e-01,  6.6251e-01,  5.4847e-01,  ..., -2.6161e-01, -1.3067e-0
1, -1.2770e-01],
          [ 6.4164e-01,  2.9999e-01,  2.8895e-01,  ..., -2.5167e-01, -2.0480e-0
1, -1.2610e-01],
          [ 5.8845e-01,  2.1644e-01,  5.9448e-01,  ...,  7.3737e-02,  1.0390e-0
1,  6.9101e-02]],

         [[-2.7824e-01, -2.6679e-01, -2.3472e-01,  ..., -2.6140e-01, -2.5957e-0
1, -2.4849e-01],
          [-2.7311e-01, -2.4047e-01, -6.5898e-02,  ..., -2.7151e-01, -2.3988e-0
1, -2.2940e-01],
          [-2.7033e-01, -2.7112e-01,  2.8378e-03,  ..., -2.6342e-01, -2.7417e-0
1, -1.6807e-01],
          ...,
          [ 2.6723e-01,  1.0032e-01,  4.2325e-01,  ..., -2.0965e-01, -2.7619e-0
1, -2.2641e-01],
          [ 3.4079e-01,  1.8129e-01,  5.0517e-01,  ..., -2.3988e-01, -1.1193e-0
1, -1.2688e-01],
          [-2.4591e-01, -2.5108e-01, -1.9036e-01,  ..., -1.5098e-01, -1.6586e-0
1, -1.5615e-01]],

         ...,

         [[ 5.1924e-02, -5.3519e-02, -1.2956e-01,  ..., -2.5755e-01, -2.0689e-0
1, -2.1633e-01],
          [-4.6771e-02,  4.6433e-02, -3.5492e-02,  ..., -2.5216e-01, -2.3168e-0
1, -2.2613e-01],
          [ 7.8561e-03, -1.7492e-02, -7.9215e-02,  ..., -2.1185e-01, -2.1273e-0
1, -2.2894e-01],
          ...,
          [-1.5160e-01, -1.7551e-01, -1.8522e-01,  ..., -2.7179e-01, -2.6663e-0
1, -2.2627e-01],
          [-2.5968e-01, -2.6013e-01, -2.7238e-01,  ..., -2.5671e-01, -2.2406e-0
1, -1.9727e-01],
          [-2.6751e-01, -2.5728e-01, -2.7129e-01,  ..., -2.6569e-01, -2.4705e-0
1, -2.2162e-01]],

         [[-2.5939e-01, -1.8062e-01, -1.8637e-01,  ..., -2.7222e-01, -2.5198e-0
1, -2.3310e-01],
```

```
          [-2.3517e-01, -1.9173e-01, -1.1079e-01,  ..., -2.6092e-01, -2.5727e-0
1, -2.6048e-01],
          [-1.9534e-01, -1.5953e-01, -1.5045e-01,  ..., -2.5833e-01, -2.1620e-0
1, -2.2886e-01],
          ...,
          [-2.4795e-01, -2.7846e-01, -2.6397e-01,  ..., -2.7427e-01, -2.7108e-0
1, -2.7837e-01],
          [-2.1822e-01, -2.6514e-01, -2.7808e-01,  ..., -2.1156e-01, -1.5706e-0
1, -1.9673e-01],
          [-2.7812e-01, -2.7613e-01, -2.7821e-01,  ..., -4.4642e-04,  1.0366e-0
1, -9.7283e-02]],

         [[-2.2856e-01, -2.4050e-01, -2.1018e-01,  ..., -2.2955e-01, -2.1842e-0
1, -1.8160e-01],
          [-1.9673e-01, -2.2976e-01, -2.7843e-01,  ..., -2.7555e-01, -2.3458e-0
1, -2.2582e-01],
          [-2.2227e-01, -2.7282e-01, -2.6572e-01,  ..., -2.7718e-01, -2.5351e-0
1, -2.5214e-01],
          ...,
          [-2.3298e-01, -2.7041e-01, -2.5213e-01,  ..., -2.3040e-01, -2.0790e-0
1, -2.5009e-01],
          [-2.3310e-01, -2.7630e-01, -2.5267e-01,  ..., -1.9440e-01, -1.7419e-0
1, -2.1922e-01],
          [-1.7990e-01, -1.6084e-01, -2.1372e-01,  ..., -2.7846e-01, -2.7070e-0
1, -2.0784e-01]]]], device='cuda:0')
```

In [71]:

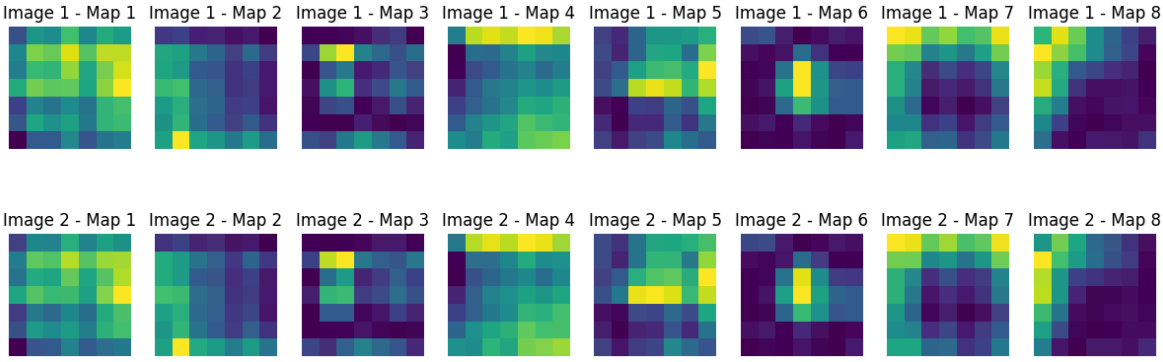
```python
# Visualize the feature map
def visualize_feature_maps(feature_maps):
    """Visualize the feature maps from a CNN backbone."""
    feature_maps = feature_maps.squeeze(0)  # Remove batch dimension (1, C, H, W
    num_channels = feature_maps.shape[0]  # Number of feature maps

    # Plot first 16 feature maps (or fewer if not available)
    num_cols = 4
    num_rows = min(4, num_channels // num_cols + 1)

    fig, axes = plt.subplots(num_rows, num_cols, figsize=(10, 10))
    axes = axes.flatten()

    for i in range(min(num_channels, 16)):  # Limit to 16 feature maps
        ax = axes[i]
        ax.imshow(feature_maps[i].cpu().detach().numpy(), cmap='viridis')
        ax.axis("off")
        ax.set_title(f"Feature Map {i+1}")

    plt.tight_layout()
    plt.show()

visualize_feature_maps(backbone(img_tensor))
```



In [72]:
```
import cv2
import torch
import numpy as np
import matplotlib.pyplot as plt
import torch.nn.functional as F
from torchvision import transforms

device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

# Move model to device
backbone = backbone.to(device)

# Function to resize and preprocess image (including BGR to RGB conversion)
def preprocess_image(img, target_size=(224, 224)):
    # Convert from BGR to RGB
    img_rgb = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

    # Convert to PIL Image, resize, and then to Tensor
    transform = transforms.Compose([
        transforms.ToPILImage(),
        transforms.Resize(target_size),
        transforms.ToTensor(),
    ])

    # Convert to tensor and move to device
    img_tensor = transform(img_rgb).unsqueeze(0).float().to(device) / 255.0  # N
```

```
    return img_tensor

# Load images
img1 = cv2.imread(img_path)
img2 = cv2.imread('/kaggle/input/cat-and-dog/training_set/training_set/dogs/dog.

# Preprocess images
img1_tensor = preprocess_image(img1)  # First image
img2_tensor = preprocess_image(img2)  # Second image

# Pass images through the backbone to extract feature maps
feature_maps1 = backbone(img1_tensor)
feature_maps2 = backbone(img2_tensor)

# Ensure both feature maps have the same shape
assert feature_maps1.shape == feature_maps2.shape, "Feature map dimensions do no

# Function to compute cosine similarity between two feature maps
def compute_similarity(fmap1, fmap2):
    fmap1 = fmap1.flatten(start_dim=1)  # Flatten each feature map to a vector
    fmap2 = fmap2.flatten(start_dim=1)
    similarity = F.cosine_similarity(fmap1, fmap2, dim=1)  # Compute cosine simi
    return similarity.mean().item()  # Return average similarity score

# Compute similarity between the feature maps
similarity_score = compute_similarity(feature_maps1, feature_maps2)
print(f"Feature Map Similarity: {similarity_score:.4f}")

# Function to visualize feature maps
def visualize_feature_maps(fmap1, fmap2):
    fmap1 = fmap1.squeeze(0).cpu().detach().numpy()  # Remove batch dimension
    fmap2 = fmap2.squeeze(0).cpu().detach().numpy()

    num_feature_maps = min(8, fmap1.shape[0])  # Show up to 8 feature maps
    fig, axes = plt.subplots(2, num_feature_maps, figsize=(15, 5))

    for i in range(num_feature_maps):
        axes[0, i].imshow(fmap1[i], cmap='viridis')
        axes[0, i].axis("off")
        axes[0, i].set_title(f"Image 1 - Map {i+1}")

        axes[1, i].imshow(fmap2[i], cmap='viridis')
        axes[1, i].axis("off")
        axes[1, i].set_title(f"Image 2 - Map {i+1}")

    plt.show()

# Visualize feature maps for comparison
visualize_feature_maps(feature_maps1, feature_maps2)
```

```
Feature Map Similarity: 0.9996
```



Custom Dataset

In [73]:
```
# Loading Dataset

from roboflow import Roboflow
rf = Roboflow(api_key="YO8Qsooc19zEzPuFRZoh")
project = rf.workspace("sunfire-marik-alar9").project("crime-cctv-object-detecti
version = project.version(8)
dataset = version.download("yolov11")
```

```
loading Roboflow workspace...
loading Roboflow project...
Downloading Dataset Version Zip in Crime-CCTV-Object-Detection-8 to yolov11:: 10
0%|██████████| 197918/197918 [00:06<00:00, 29274.40it/s]
Extracting Dataset Version Zip to Crime-CCTV-Object-Detection-8 in yolov11:: 100%
|██████████| 7038/7038 [00:01<00:00, 6993.70it/s]
```

In [74]:
```
# Training Model
history = model.train(data='/kaggle/working/Crime-CCTV-Object-Detection-8/data.y
```

```
Ultralytics 8.3.78 🚀 Python-3.10.12 torch-2.5.1+cu121 CUDA:0 (Tesla P100-PCIE-1
6GB, 16269MiB)
engine/trainer: task=detect, mode=train, model=yolo11n.pt, data=/kaggle/working/C
rime-CCTV-Object-Detection-8/data.yaml, epochs=1, time=None, patience=100, batch=
16, imgsz=128, save=True, save_period=-1, cache=False, device=None, workers=8, pr
oject=None, name=train2, exist_ok=False, pretrained=True, optimizer=auto, verbose
=True, seed=0, deterministic=True, single_cls=False, rect=False, cos_lr=False, cl
ose_mosaic=10, resume=False, amp=True, fraction=1.0, profile=False, freeze=None,
multi_scale=False, overlap_mask=True, mask_ratio=4, dropout=0.0, val=True, split=
val, save_json=False, save_hybrid=False, conf=None, iou=0.7, max_det=300, half=Fa
lse, dnn=False, plots=True, source=None, vid_stride=1, stream_buffer=False, visua
lize=False, augment=False, agnostic_nms=False, classes=None, retina_masks=False,
embed=None, show=False, save_frames=False, save_txt=False, save_conf=False, save_
crop=False, show_labels=True, show_conf=True, show_boxes=True, line_width=None, f
ormat=torchscript, keras=False, optimize=False, int8=False, dynamic=False, simpli
fy=True, opset=None, workspace=None, nms=False, lr0=0.01, lrf=0.01, momentum=0.93
7, weight_decay=0.0005, warmup_epochs=3.0, warmup_momentum=0.8, warmup_bias_lr=0.
1, box=7.5, cls=0.5, dfl=1.5, pose=12.0, kobj=1.0, nbs=64, hsv_h=0.015, hsv_s=0.
7, hsv_v=0.4, degrees=0.0, translate=0.1, scale=0.5, shear=0.0, perspective=0.0,
flipud=0.0, fliplr=0.5, bgr=0.0, mosaic=1.0, mixup=0.0, copy_paste=0.0, copy_past
e_mode=flip, auto_augment=randaugment, erasing=0.4, crop_fraction=1.0, cfg=None,
tracker=botsort.yaml, save_dir=runs/detect/train2
Overriding model.yaml nc=80 with nc=3

                   from  n    params  module
arguments
  0                  -1  1       464  ultralytics.nn.modules.conv.Conv
[3, 16, 3, 2]
  1                  -1  1      4672  ultralytics.nn.modules.conv.Conv
[16, 32, 3, 2]
  2                  -1  1      6640  ultralytics.nn.modules.block.C3k2
[32, 64, 1, False, 0.25]
  3                  -1  1     36992  ultralytics.nn.modules.conv.Conv
[64, 64, 3, 2]
  4                  -1  1     26080  ultralytics.nn.modules.block.C3k2
[64, 128, 1, False, 0.25]
  5                  -1  1    147712  ultralytics.nn.modules.conv.Conv
[128, 128, 3, 2]
  6                  -1  1     87040  ultralytics.nn.modules.block.C3k2
[128, 128, 1, True]
  7                  -1  1    295424  ultralytics.nn.modules.conv.Conv
[128, 256, 3, 2]
  8                  -1  1    346112  ultralytics.nn.modules.block.C3k2
[256, 256, 1, True]
  9                  -1  1    164608  ultralytics.nn.modules.block.SPPF
[256, 256, 5]
 10                  -1  1    249728  ultralytics.nn.modules.block.C2PSA
[256, 256, 1]
 11                  -1  1         0  torch.nn.modules.upsampling.Upsample
[None, 2, 'nearest']
 12             [-1, 6]  1         0  ultralytics.nn.modules.conv.Concat
[1]
 13                  -1  1    111296  ultralytics.nn.modules.block.C3k2
[384, 128, 1, False]
 14                  -1  1         0  torch.nn.modules.upsampling.Upsample
[None, 2, 'nearest']
 15             [-1, 4]  1         0  ultralytics.nn.modules.conv.Concat
[1]
 16                  -1  1     32096  ultralytics.nn.modules.block.C3k2
[256, 64, 1, False]
```

```
 17                  -1  1     36992  ultralytics.nn.modules.conv.Conv
[64, 64, 3, 2]
 18            [-1, 13]  1         0  ultralytics.nn.modules.conv.Concat
[1]
 19                  -1  1     86720  ultralytics.nn.modules.block.C3k2
[192, 128, 1, False]
 20                  -1  1    147712  ultralytics.nn.modules.conv.Conv
[128, 128, 3, 2]
 21            [-1, 10]  1         0  ultralytics.nn.modules.conv.Concat
[1]
 22                  -1  1    378880  ultralytics.nn.modules.block.C3k2
[384, 256, 1, True]
 23        [16, 19, 22]  1    431257  ultralytics.nn.modules.head.Detect
[3, [64, 128, 256]]
YOLO11n summary: 181 layers, 2,590,425 parameters, 2,590,409 gradients, 6.4 GFLOP
s

Transferred 79/499 items from pretrained weights
TensorBoard: Start with 'tensorboard --logdir runs/detect/train2', view at htt
p://localhost:6006/
Freezing layer 'model.23.dfl.conv.weight'
AMP: running Automatic Mixed Precision (AMP) checks...
AMP: checks passed ✅
train: Scanning /kaggle/working/Crime-CCTV-Object-Detection-8/train/labels... 291
3 images, 0 backgrounds, 0 corrupt: 100%|██████████| 2913/2913 [00:02<00:00, 110
6.71it/s]
train: New cache created: /kaggle/working/Crime-CCTV-Object-Detection-8/train/lab
els.cache
albumentations: Blur(p=0.01, blur_limit=(3, 7)), MedianBlur(p=0.01, blur_limit=
(3, 7)), ToGray(p=0.01, num_output_channels=3, method='weighted_average'), CLAHE
(p=0.01, clip_limit=(1.0, 4.0), tile_grid_size=(8, 8))
val: Scanning /kaggle/working/Crime-CCTV-Object-Detection-8/valid/labels... 400 i
mages, 0 backgrounds, 0 corrupt: 100%|██████████| 400/400 [00:00<00:00, 773.10it/
s]
val: New cache created: /kaggle/working/Crime-CCTV-Object-Detection-8/valid/label
s.cache
Plotting labels to runs/detect/train2/labels.jpg...
optimizer: 'optimizer=auto' found, ignoring 'lr0=0.01' and 'momentum=0.937' and d
etermining best 'optimizer', 'lr0' and 'momentum' automatically...
optimizer: AdamW(lr=0.001429, momentum=0.9) with parameter groups 81 weight(decay
=0.0), 88 weight(decay=0.0005), 87 bias(decay=0.0)
TensorBoard: model graph visualization added ✅
Image sizes 128 train, 128 val
Using 4 dataloader workers
Logging results to runs/detect/train2
Starting training for 1 epochs...

      Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
        1/1     0.422G      3.047      3.532      2.348          2        128: 10
0%|██████████| 183/183 [00:19<00:00,  9.23it/s]
                 Class     Images  Instances      Box(P          R      mAP50  mA
P50-95): 100%|██████████| 13/13 [00:02<00:00,  5.59it/s]
                   all        400        510       0.12      0.206     0.0889
0.0329
invalid value encountered in less
invalid value encountered in less
```

```
1 epochs completed in 0.009 hours.
Optimizer stripped from runs/detect/train2/weights/last.pt, 5.4MB
Optimizer stripped from runs/detect/train2/weights/best.pt, 5.4MB

Validating runs/detect/train2/weights/best.pt...
Ultralytics 8.3.78 🚀 Python-3.10.12 torch-2.5.1+cu121 CUDA:0 (Tesla P100-PCIE-16GB, 16269MiB)
YOLO11n summary (fused): 100 layers, 2,582,737 parameters, 0 gradients, 6.3 GFLOPs
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 13/13 [00:01<00:00,  6.52it/s]
                   all        400        510       0.12      0.212     0.0886     0.0328
              burglary        142        170     0.0979        0.2     0.0729     0.0249
              fighting        135        159      0.188      0.308      0.157     0.0601
               robbery        123        181     0.0742      0.127     0.0356     0.0133
invalid value encountered in less
invalid value encountered in less
Speed: 0.0ms preprocess, 1.6ms inference, 0.0ms loss, 0.8ms postprocess per image
Results saved to runs/detect/train2
```

In [75]:
```
# Predicting Model
results = model.predict('/kaggle/working/Crime-CCTV-Object-Detection-8/train/ima
results[0].show()
```

```
image 1/1 /kaggle/working/Crime-CCTV-Object-Detection-8/train/images/ROb10_0064_jpg.rf.74558bdb832fd2b9a0f1f064b4882fe1.jpg: 128x128 (no detections), 11.5ms
Speed: 0.6ms preprocess, 11.5ms inference, 0.6ms postprocess per image at shape (1, 3, 128, 128)
```

Our YOLO Class

In [77]:
```
# Our Class
class YOLO_DSC:
    def __init__(self, model_path='yolo11n.pt', threshold = 0.434):
        self.model = YOLO(model_path)
        self.threshold = threshold
        self.backbone = self.model.model.model[:10]

    def preprocess_img(self, img_path):
        img = cv2.cvtColor(cv2.imread(img_path), cv2.COLOR_BGR2RGB)
        return img

    def get_featuremap(self, img_path):
        img = self.preprocess_img(img_path)
        results = self.model.predict(img)

        feature_maps = {}


        for result in results:
            boxes = result.boxes.xyxy
            conf = result.boxes.conf
```

```
            for i in range(len(boxes)):
                if conf[i] > self.threshold:
                    x1, y1, x2, y2 = map(int, boxes[i])
                    obj_crop = img[y1:y2, x1:x2]

                    obj_tensor = (torch.tensor(obj_crop).permute(2, 0, 1).float(
                    feature_maps[i] = self.backbone(obj_tensor)
        return feature_maps

    def train_(data, epochs, imgsz):
        history = self.model.train(data=data, epochs=epochs, imgsz=imgsz)
        return history
```

In [78]:
```
obj = YOLO_DSC()
```

In [79]:
```
obj.get_featuremap(img_path)
```

```
0: 480x640 3 cats, 1 bowl, 12.2ms
Speed: 2.4ms preprocess, 12.2ms inference, 1.5ms postprocess per image at shape
(1, 3, 480, 640)
```

Out[79]:
```
{0: tensor([[[[-0.2423, -0.2534, -0.2233,  ..., -0.1753, -0.2583, -0.1938],
              [-0.2757, -0.2615, -0.2392,  ..., -0.1759, -0.2380, -0.1883],
              [-0.2601, -0.2356, -0.1823,  ..., -0.2152, -0.2631, -0.2151],
              [-0.2637, -0.2386, -0.1774,  ..., -0.2196, -0.2535, -0.2727],
              [-0.2309, -0.2579, -0.1845,  ..., -0.2169, -0.2782, -0.2741],
              [-0.1898, -0.2531, -0.1966,  ..., -0.2355, -0.2488, -0.1813]],

             [[ 0.7820,  0.8527,  0.3970,  ...,  0.1815,  0.3247,  0.0644],
              [ 0.4854,  0.5805,  0.2971,  ...,  0.3340,  0.0036,  0.0188],
              [ 0.5573,  0.6076,  0.3849,  ...,  0.0924, -0.0310, -0.1659],
              [ 0.8334,  0.5030,  0.1906,  ...,  0.1006,  0.2305,  0.0685],
              [ 0.5516,  0.2925,  0.1505,  ..., -0.0978, -0.0949, -0.0454],
              [ 0.7342,  0.6466,  0.5241,  ...,  0.1091,  0.2977,  0.2201]],

             [[-0.2078, -0.2643, -0.2785,  ..., -0.2774, -0.2756, -0.2444],
              [-0.2678, -0.2674, -0.2655,  ..., -0.2784, -0.2755, -0.2174],
              [-0.2730, -0.2420, -0.0542,  ..., -0.2233, -0.2524, -0.2678],
              [-0.2624,  0.0126, -0.0539,  ..., -0.2691, -0.2782, -0.2351],
              [-0.1389, -0.0713,  0.2913,  ..., -0.2784, -0.2740, -0.2761],
              [-0.2621, -0.2784, -0.2744,  ..., -0.2095, -0.2128, -0.2167]],

             ...,

             [[-0.2037, -0.1859, -0.1250,  ..., -0.2734, -0.1216, -0.0365],
              [-0.1768, -0.1675, -0.0252,  ..., -0.2711, -0.2746, -0.2596],
              [-0.1615, -0.1885, -0.0958,  ..., -0.2382, -0.2785, -0.2784],
              [-0.2278, -0.1992, -0.1723,  ..., -0.1061, -0.2724, -0.2765],
              [-0.2695, -0.2726, -0.2679,  ..., -0.2411, -0.2492, -0.2718],
              [-0.2213, -0.2134, -0.2375,  ..., -0.2568, -0.2766, -0.2717]],

             [[-0.1773, -0.1436, -0.1658,  ..., -0.1526, -0.1257, -0.1326],
              [-0.1497, -0.1959, -0.2708,  ..., -0.1628, -0.1818, -0.1718],
              [-0.1896, -0.1914, -0.2135,  ..., -0.2774, -0.2758, -0.2123],
              [-0.2378, -0.2475, -0.2559,  ..., -0.2596, -0.1906, -0.2505],
              [-0.2459, -0.2701, -0.2696,  ..., -0.2383, -0.2474, -0.2703],
              [-0.2457, -0.2443, -0.2529,  ..., -0.2729, -0.2478, -0.2740]],

             [[-0.2111, -0.2036, -0.1774,  ..., -0.1240, -0.1353, -0.1270],
              [-0.2103, -0.1990, -0.2047,  ..., -0.1319, -0.1397, -0.1981],
              [-0.2598, -0.2360, -0.2631,  ..., -0.1562, -0.1282, -0.2044],
              [-0.2181, -0.1918, -0.2476,  ..., -0.1656, -0.1564, -0.1683],
              [-0.2363, -0.2372, -0.2294,  ..., -0.2247, -0.1725, -0.1789],
              [-0.1689, -0.1620, -0.2202,  ..., -0.1755, -0.1446, -0.1845]]]], dev
ice='cuda:0')}
```