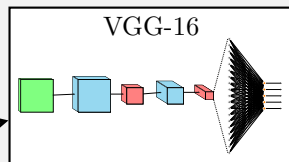


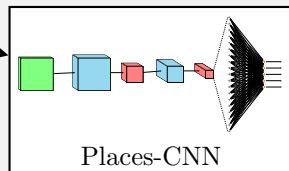


what are the *objects*
in the image?

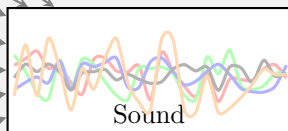


dining table ($p = 0.41$)
soup bowl ($p = 0.12$)
...

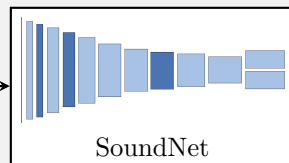
what is the *place*
in the image?



food court ($p = 0.69$)
cafeteria ($p = 0.16$)
...



what are the *objects*
and the *place*
in the sound?



play sound with min. KL

KL

KL