Architect:

fusion\_model(

(model\_feat): model\_features(

(L1): Linear(in\_features=9, out\_features=6, bias=True)

(L2): Linear(in\_features=6, out\_features=3, bias=True)

)

(model\_img): model\_image(

(L1): Conv2d(3, 64, kernel\_size=(16, 16), stride=(8, 8))

(L2): Conv2d(64, 64, kernel\_size=(4, 4), stride=(2, 2))

(L22): Conv2d(64, 64, kernel\_size=(4, 4), stride=(2, 2))

(L3): Linear(in\_features=1600, out\_features=30, bias=True)

)

(decoder): decoder(

(L1): Linear(in\_features=33, out\_features=16, bias=True)

(L2): Linear(in\_features=16, out\_features=8, bias=True)

(L3): Linear(in\_features=8, out\_features=1, bias=True)

)

)