

# Results on Test Set of Non-salient Task

*Text Input*

“““

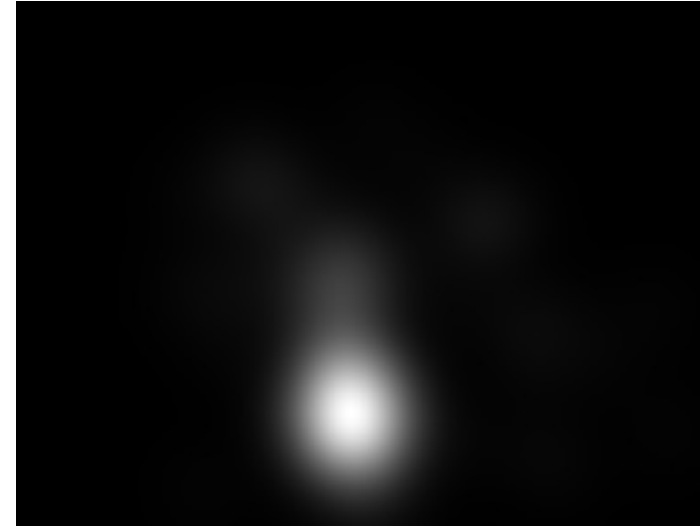
There's a guy on the road wearing  
a yellow helmet and a fluorescent  
green suit.

”””

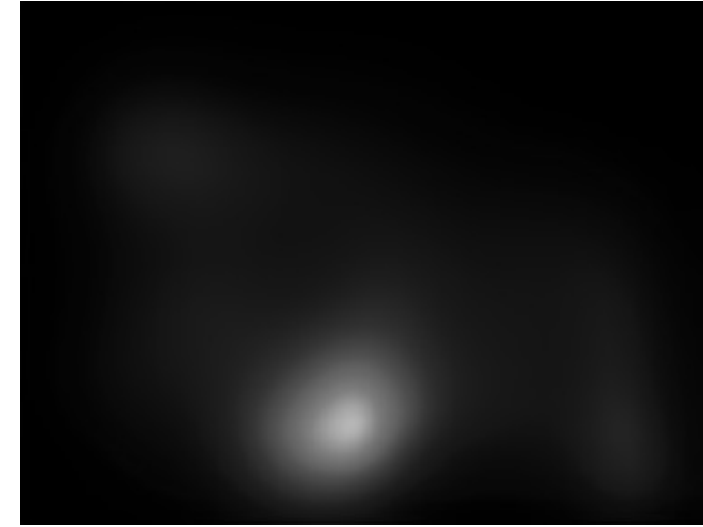
*Original image*



*Ground Truth*



*Prediction*



“““

The ship bears the symbol of a  
hexagram

”””

