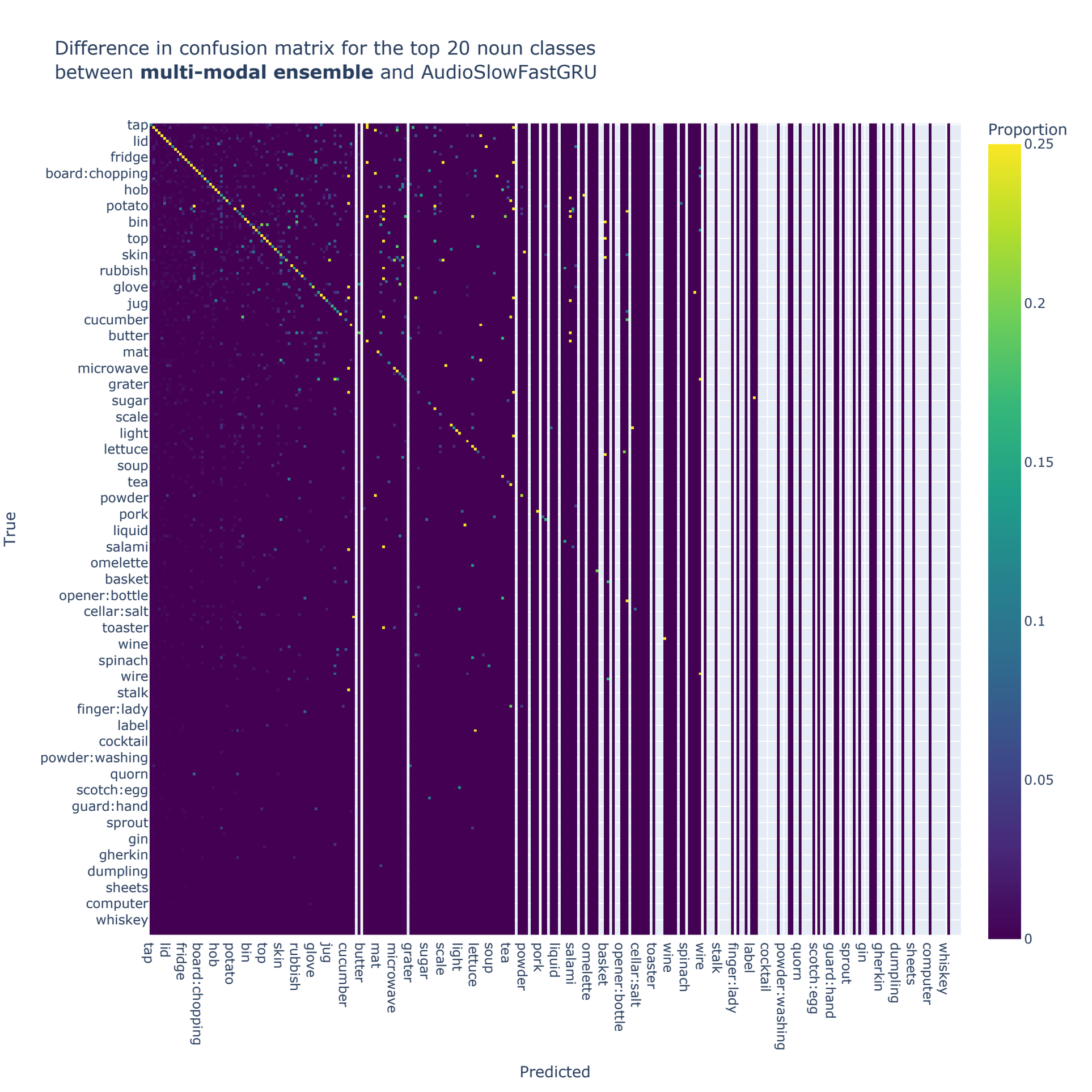
Difference in confusion matrix for the top 20 noun classes
between **multi-modal ensemble** and AudioSlowFastGRU
True
tap
lid
fridge
board:chopping
hob
potato
bin
top
skin
rubbish
glove
jug
cucumber
butter
mat
microwave
grater
sugar
scale
light
lettuce
soup
tea
powder
pork
liquid
salami
omelette
basket
opener:bottle
cellar:salt
toaster
wine
spinach
wire
stalk
finger:lady
label
cocktail
powder:washing
quorn
scotch:egg
guard:hand
sprout
gin
gherkin
dumpling
sheets
computer
whiskey
Predicted
Proportion
0.25
0.2
0.15
0.1
0.05
0