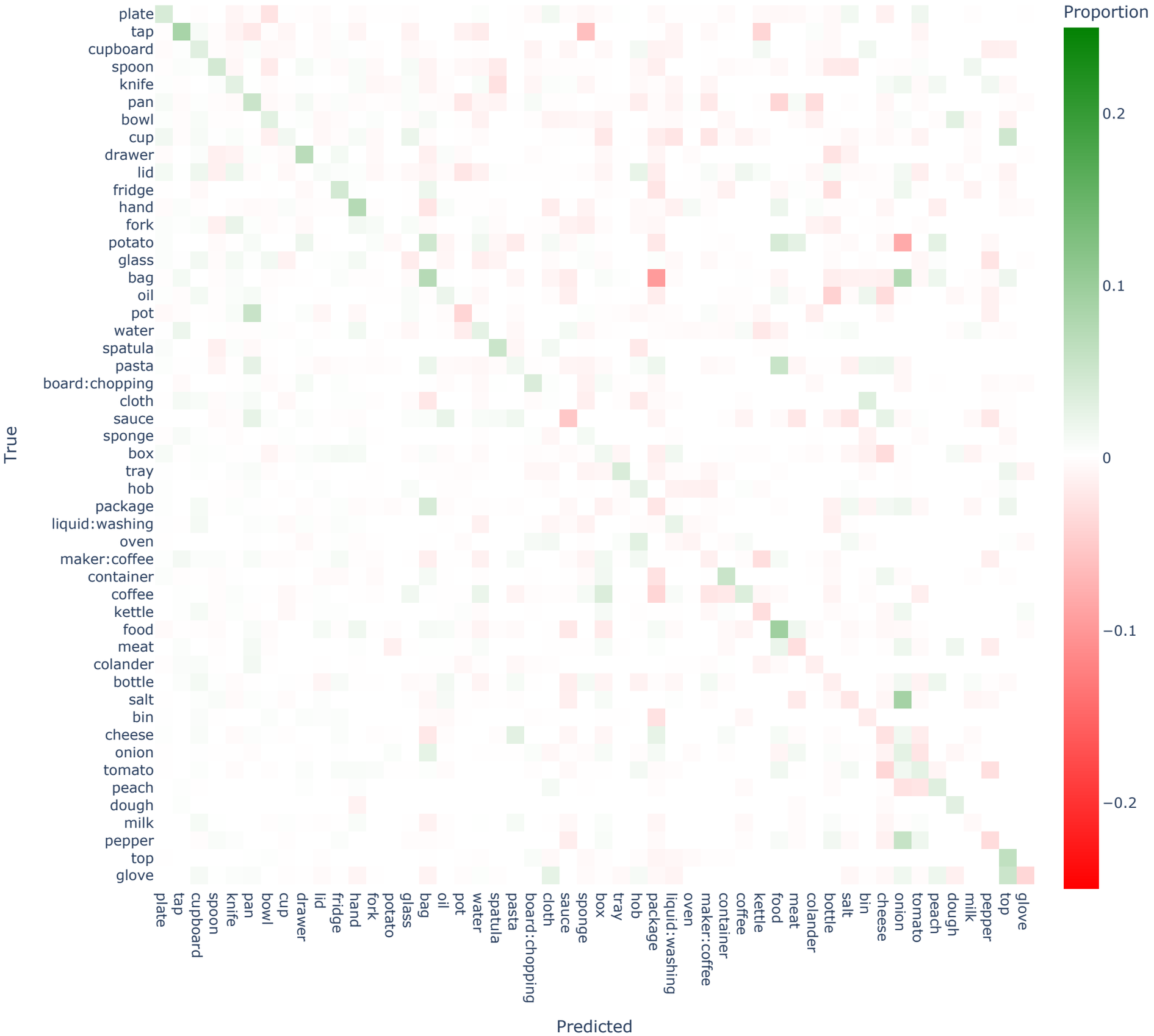Difference in confusion matrix for the top 50 noun classes between **multi-modal ensemble** and SlowFast