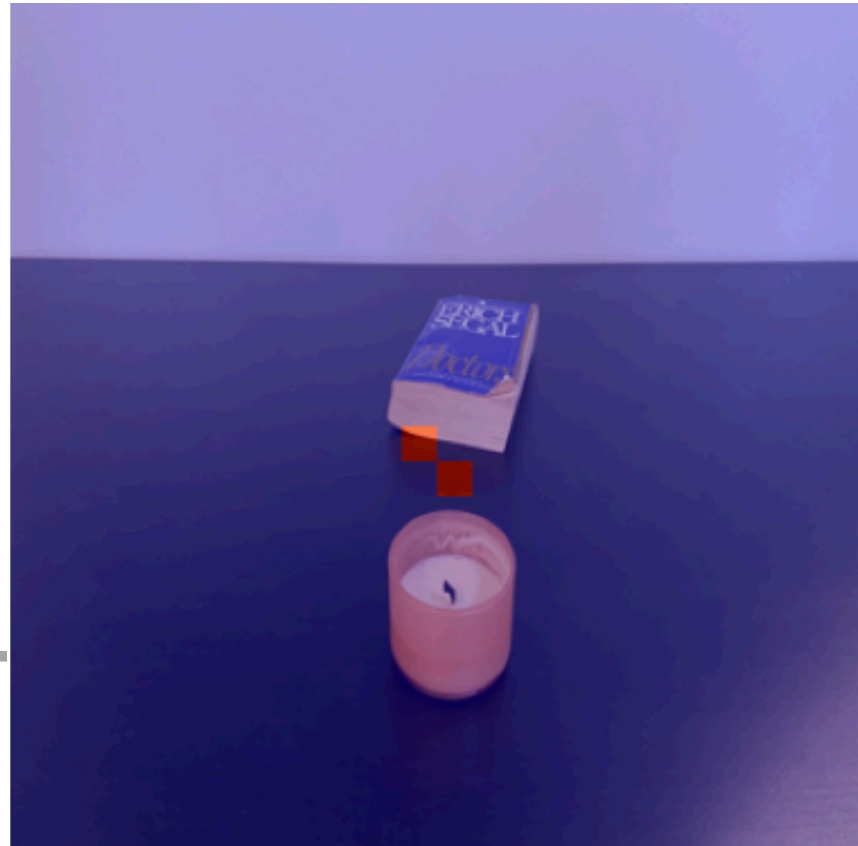


Where is the book in relation to the candle?



VLM
(e.g. LLaVA)

Attention on image



Unfamiliar!

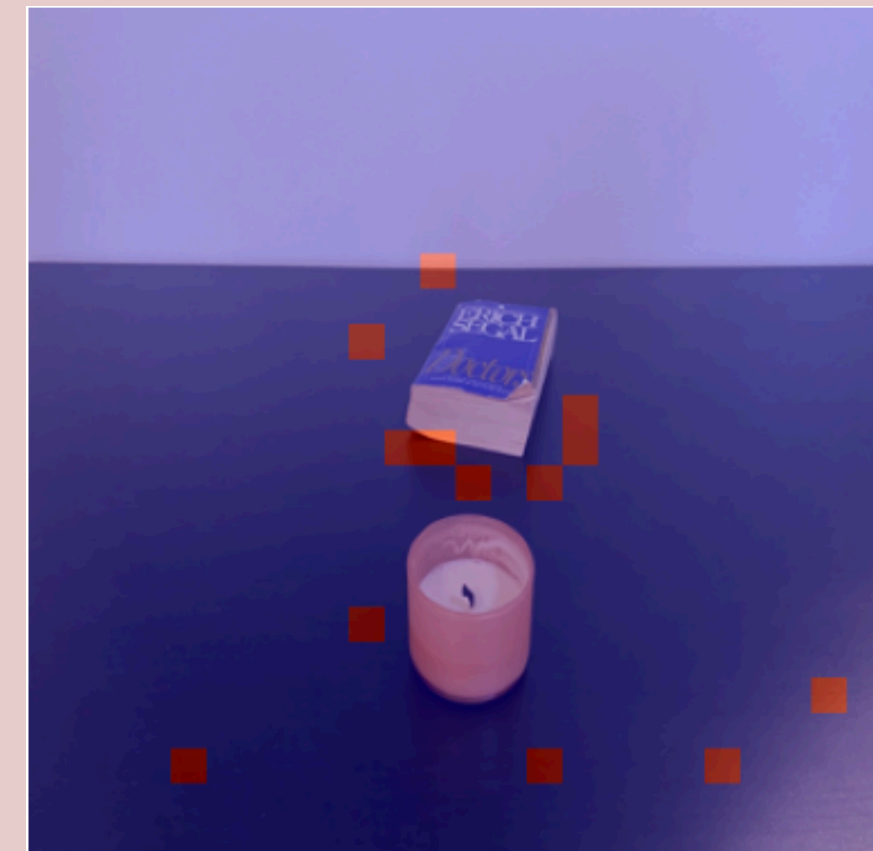
Left



Confidence: 0.27

Smoother!

Look at other patches!



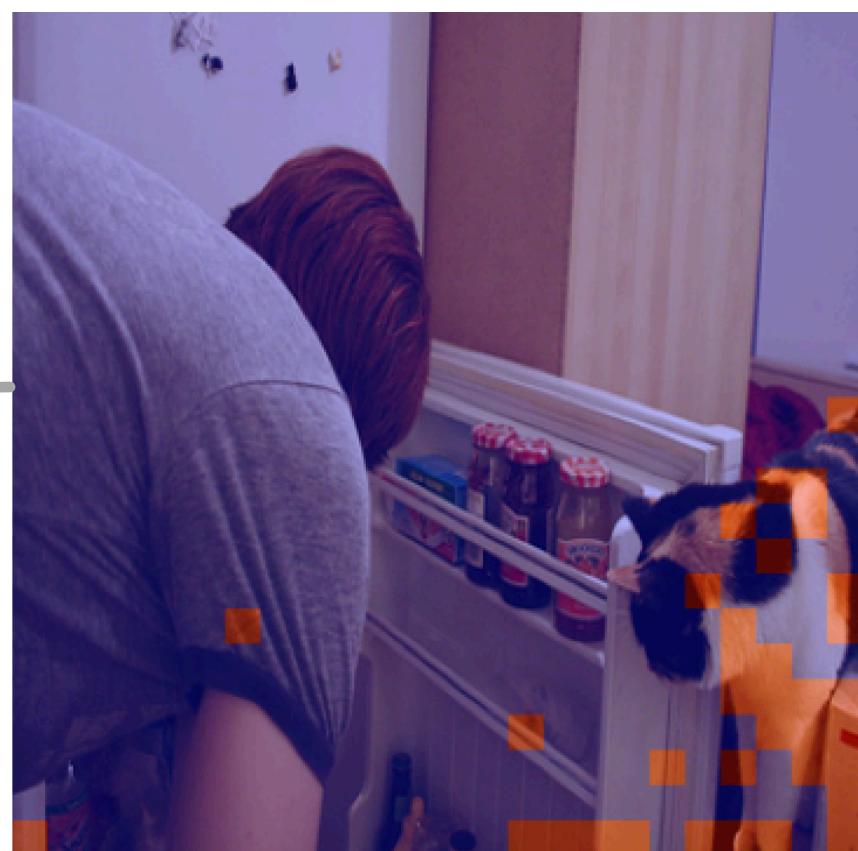
Behind



Where is the cat in relation to the person?

VLM
(e.g. LLaVA)

Attention on image



Above



Confidence: 0.38

Familiar!

Sharper!

Look more closely!



Right

