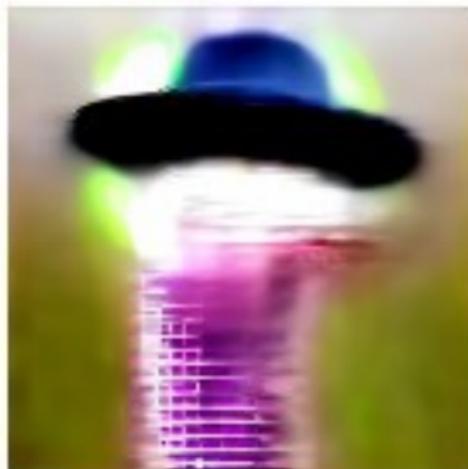
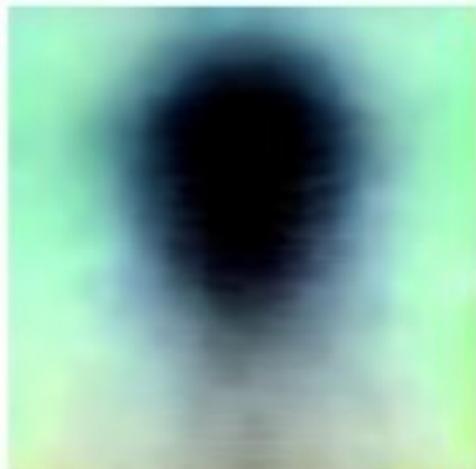




**experience**

<b>Method</b>	<b>Backbone NFE</b>	<b>Head NFE</b>	<b>Time (sec)</b>	<b>CLIPScore</b>	<b>PickScore</b>
FM	128	-	10.8	26.0	21.0
<b>DTM (our)</b>	16	8	1.6	26.8	21.21





1



12

2

12

3



12

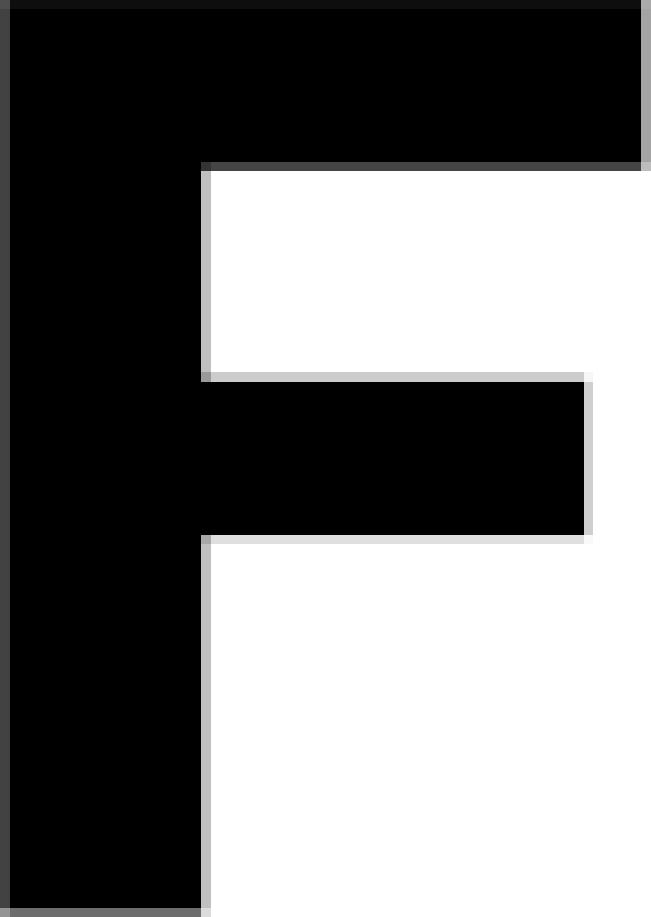
12

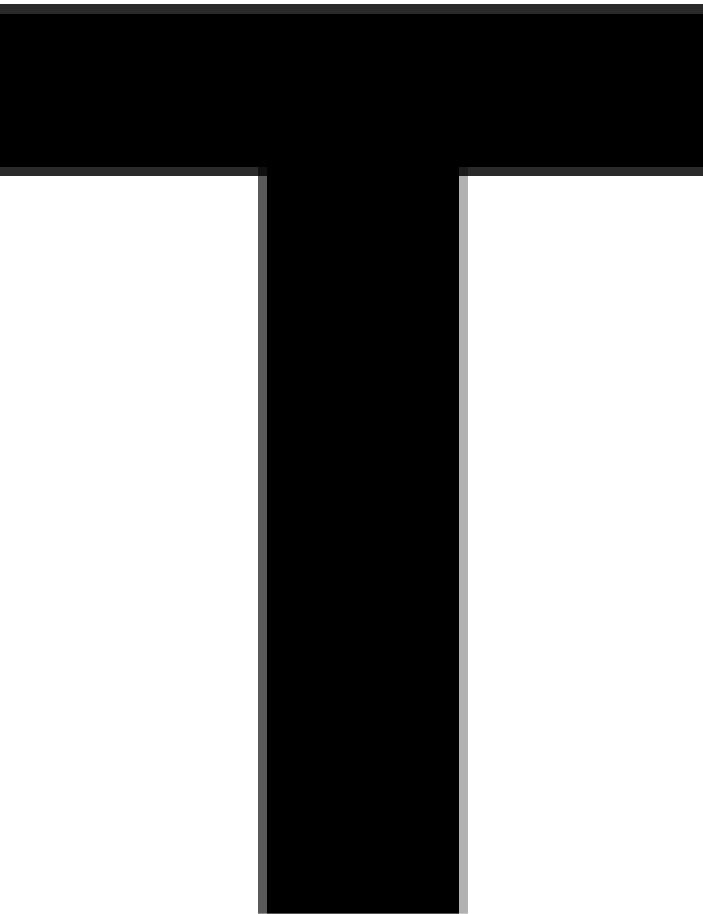
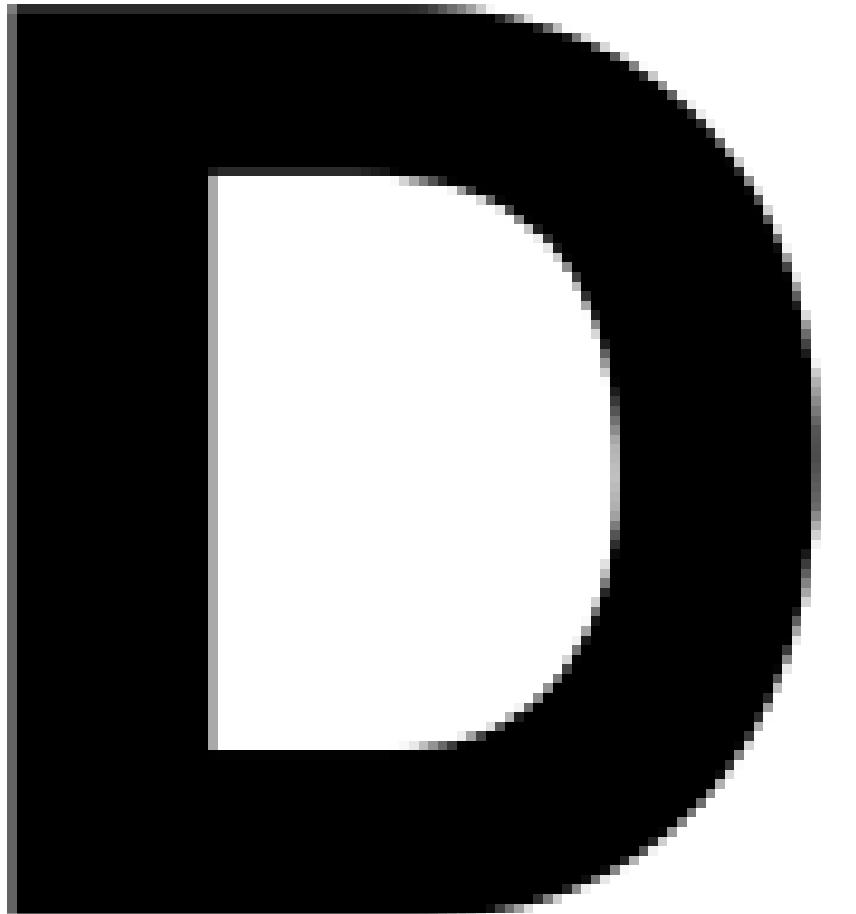


12

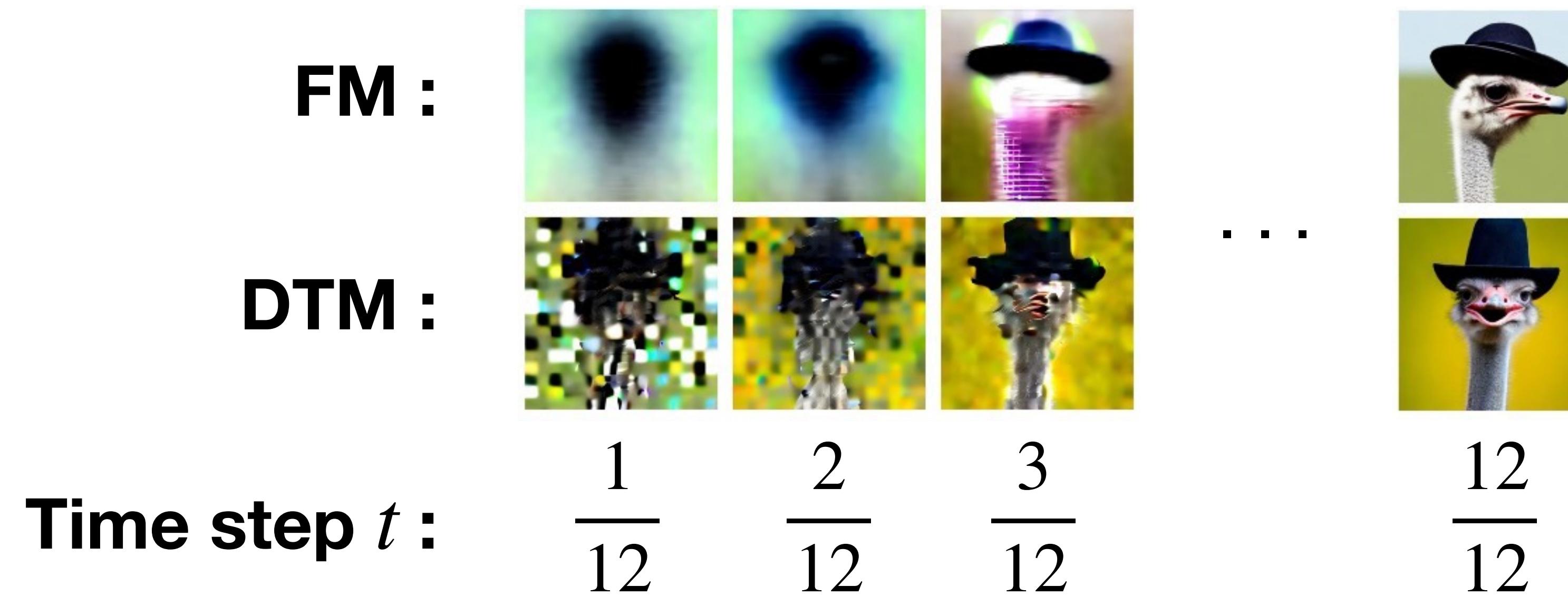
time step.







# Experiments



Method	Backbone NFE	Head NFE	Time (sec)	CLIPScore	PickScore
FM	128	-	10.8	26.0	21.0
DTM (our)	16	8	1.6	26.8	21.21

# Conclusions

1. Identify the key design choices.
2. Derive two variants: DTM and ARTM.
3. Describe relation to Flow matching.
4. Empirical validation of the method.