

Your 598 ID:

**Title of paper:**

**For all papers:**

What is their primary result?

Why is this important?

What are their key ideas? (I.e., architecture, method of training, new mathematical technique, existing technique applied in new way, something else?)

How does this work compare to prior work?

What are the limitations, either in performance or applicability?

What might be an interesting next step based on this work?

**For implementation/application papers:**

What's the architecture?

How did they train and evaluate it? (I.e., anything special about the data pipeline, loss function, optimizer?)

**For theory papers:**

Did they implement something? If so, how does the practice compare to theory?

Grader's 598 ID: