There have been a few recent methods proposed in text to video moment retrieval using natural language queries, but requiring full supervision during training.