

Soumya Jahagirdar, Minesh Mathew, Dimosthenis Karatzas, & CV Jawahar. (2023). Understanding Video Scenes Through Text: Insights from Text-Based Video Question Answering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*.