LinkedIn
ashishpatel2604

Week-2

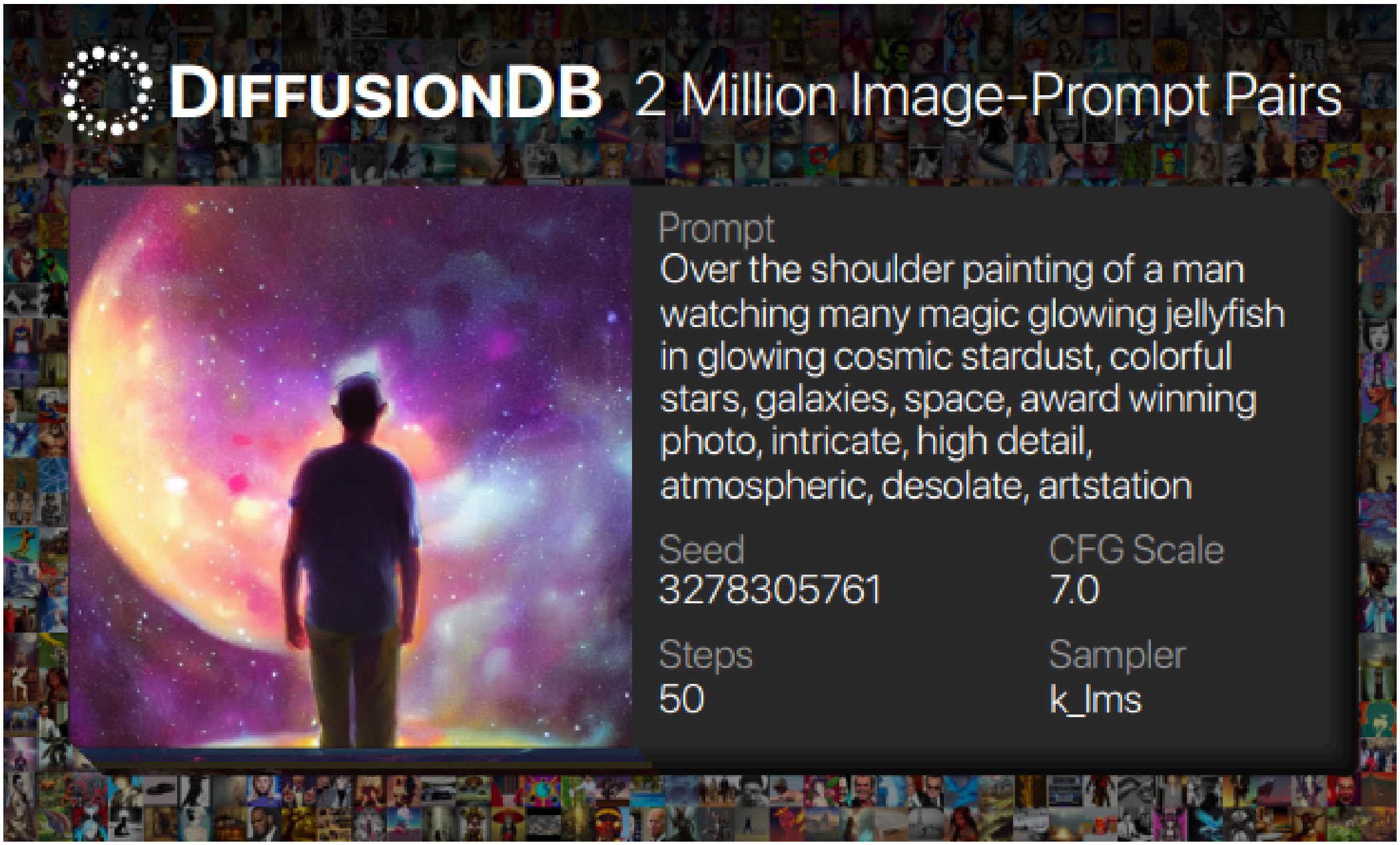
COMPUTER VISION RESEARCH PAPERS OF THE WEEK 2022

Computer Vision Group(CVG)

Follow Share

DIFFUSIONDB

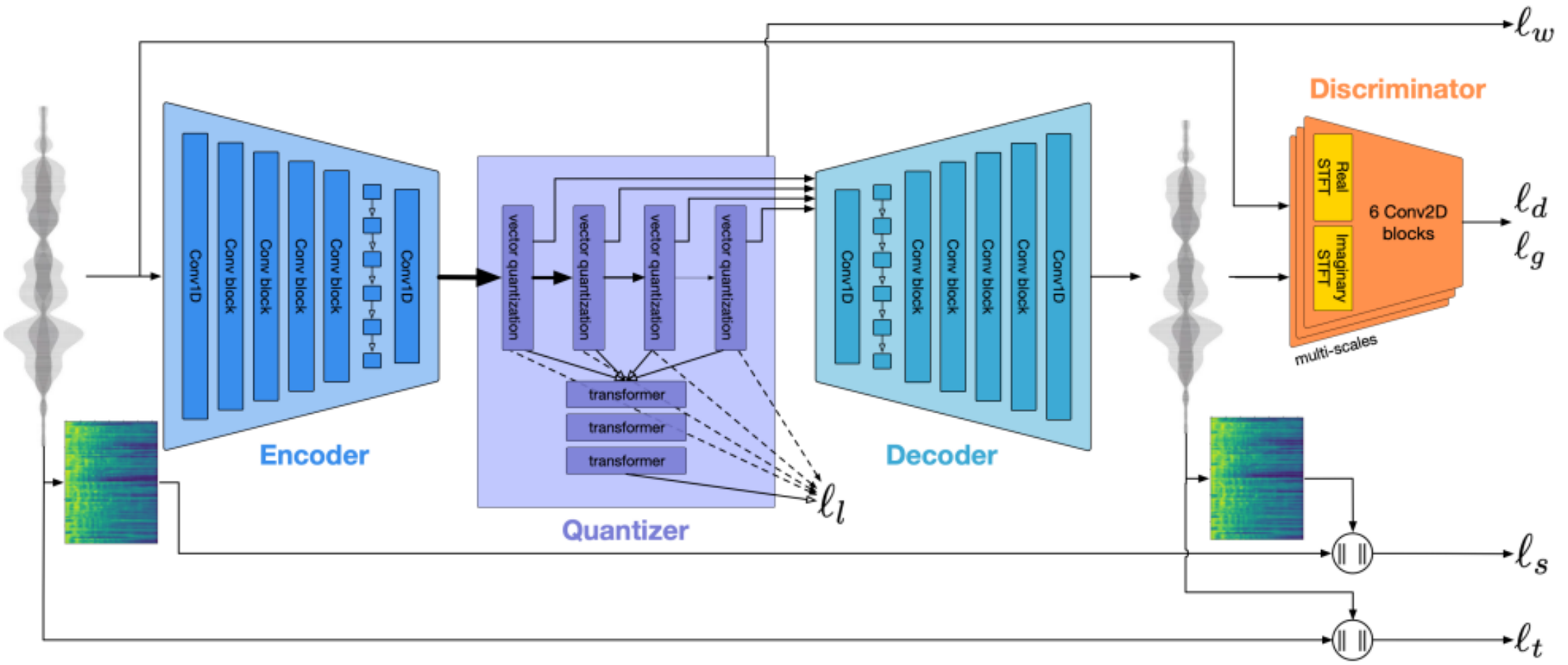
COMPUTER VISION RESEARCH PAPERS OF THE WEEK 2022



DIFFUSIONDB is the first large-scale dataset containing 2 million Stable Diffusion images and their text prompts and hyperparameters. This dataset provides exciting research opportunities in prompt engineering, deepfake detection, as well as understanding and debugging large text-to-image generative models.

ENCODEC

COMPUTER VISION RESEARCH PAPERS OF THE WEEK 2022



EnCodec : a state-of-the-art real-time neural audio compression model, producing high-fidelity audio samples across a range of sample rates and bandwidth.

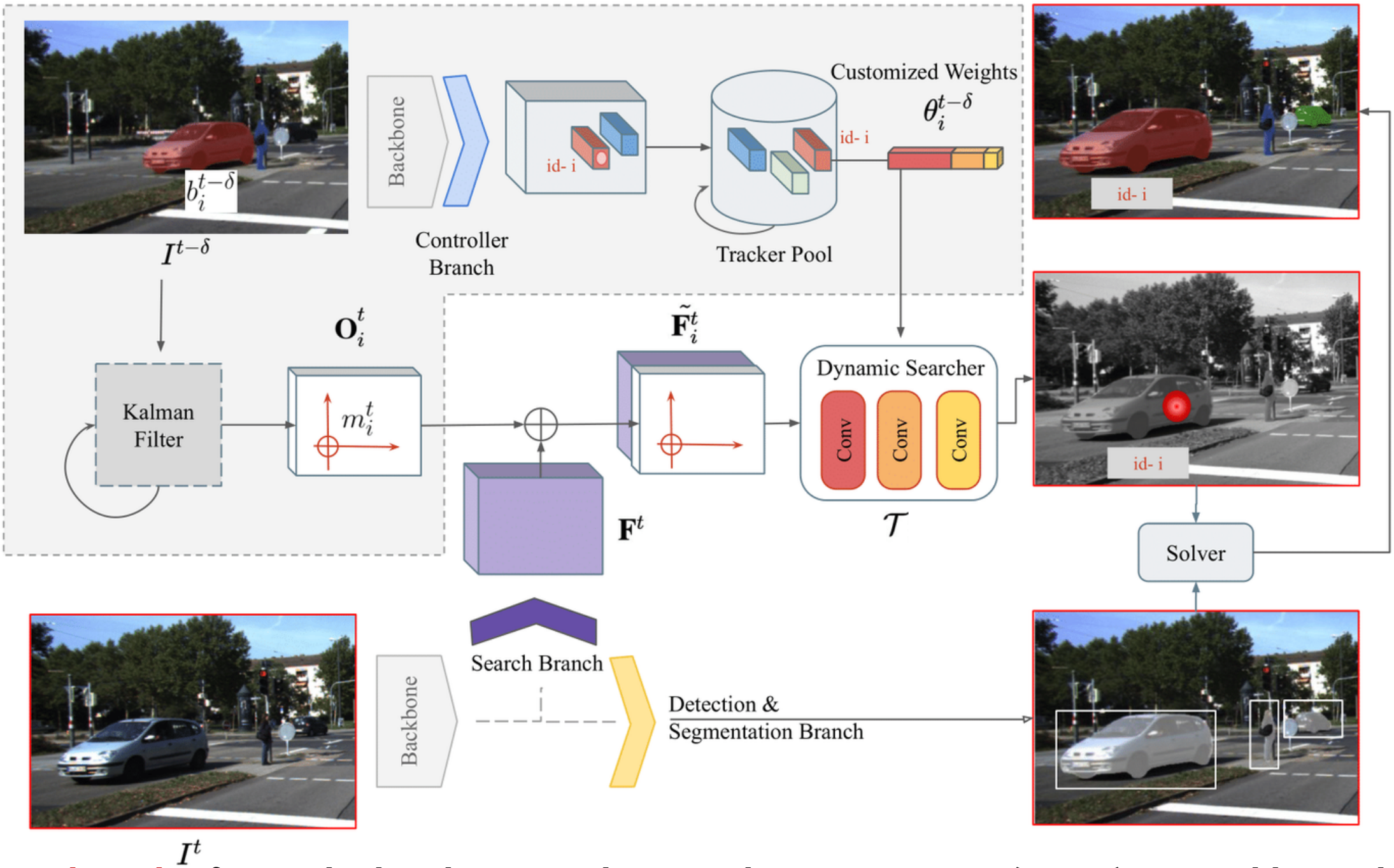
DREAMBOOTH

COMPUTER VISION RESEARCH PAPERS OF THE WEEK 2022

Google DreamBooth—our AI-powered photo booth—can generate a myriad of images of the subject in different contexts (right), using the guidance of a text prompt. The results exhibit natural interactions with the environment, as well as novel articulations and variation in lighting conditions, all while maintaining high fidelity to the key visual features of the subject.

SEARCH TRACK

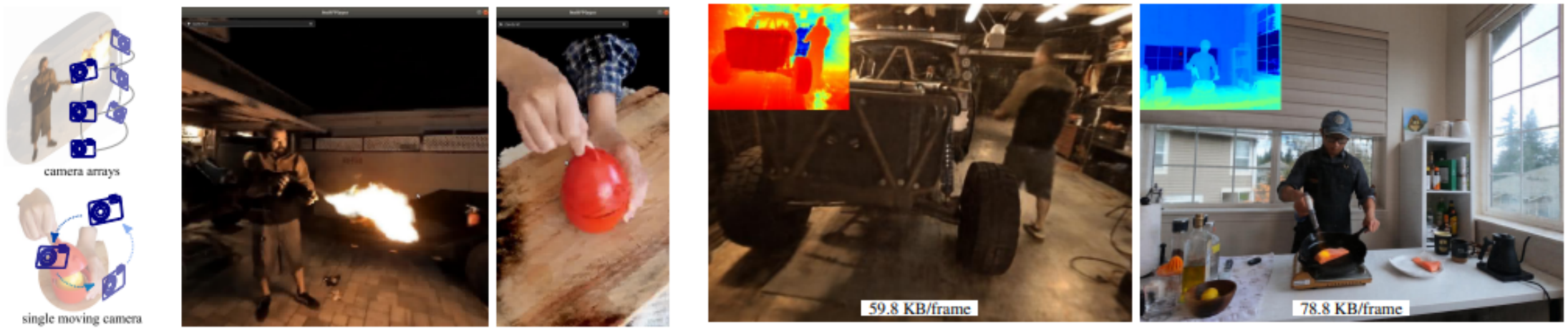
COMPUTER VISION RESEARCH PAPERS OF THE WEEK 2022



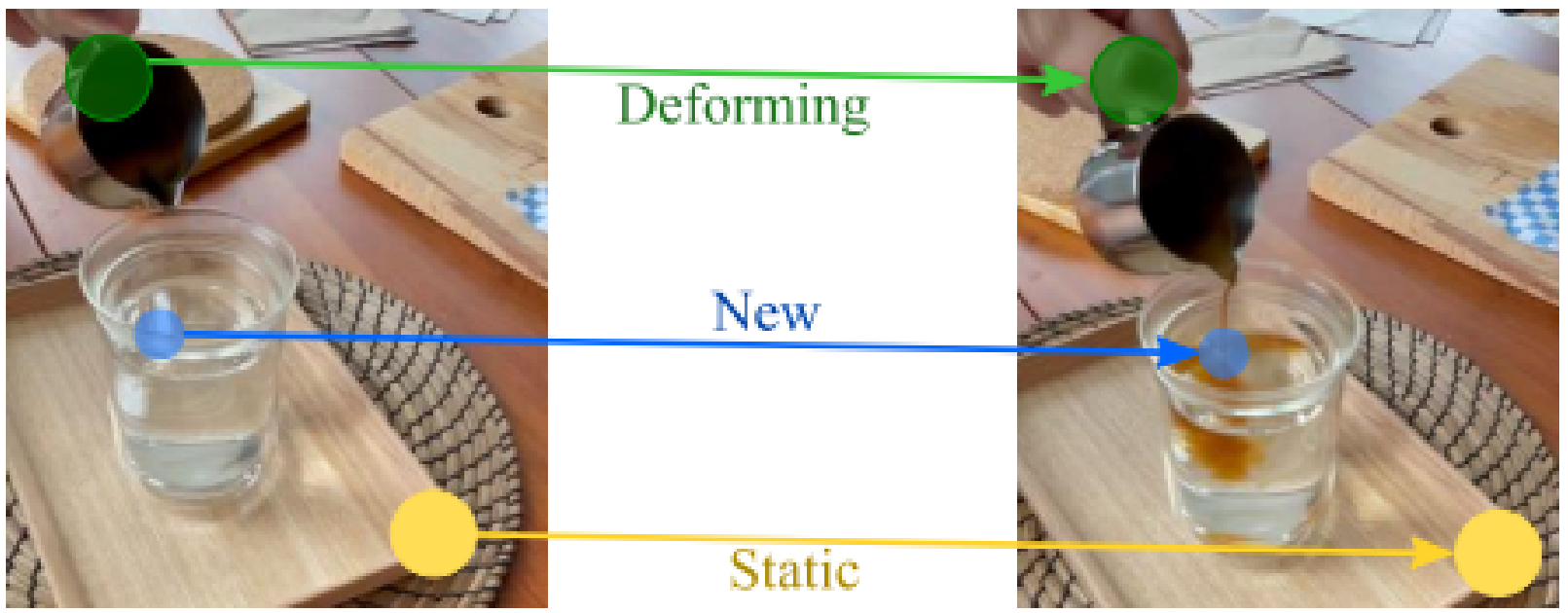
SearchTrack, for multiple object tracking and segmentation (MOTS). To address the association problem between detected objects, SearchTrack proposes object-customized search and motion-aware features. By maintaining a Kalman filter for each object, we encode the predicted motion into the motion-aware feature, which includes both motion and appearance cues. For each object, a customized fully convolutional search engine is created by SearchTrack by learning a set of weights for dynamic convolutions specific to the object.

NERFPLAYER

COMPUTER VISION RESEARCH PAPERS OF THE WEEK 2022



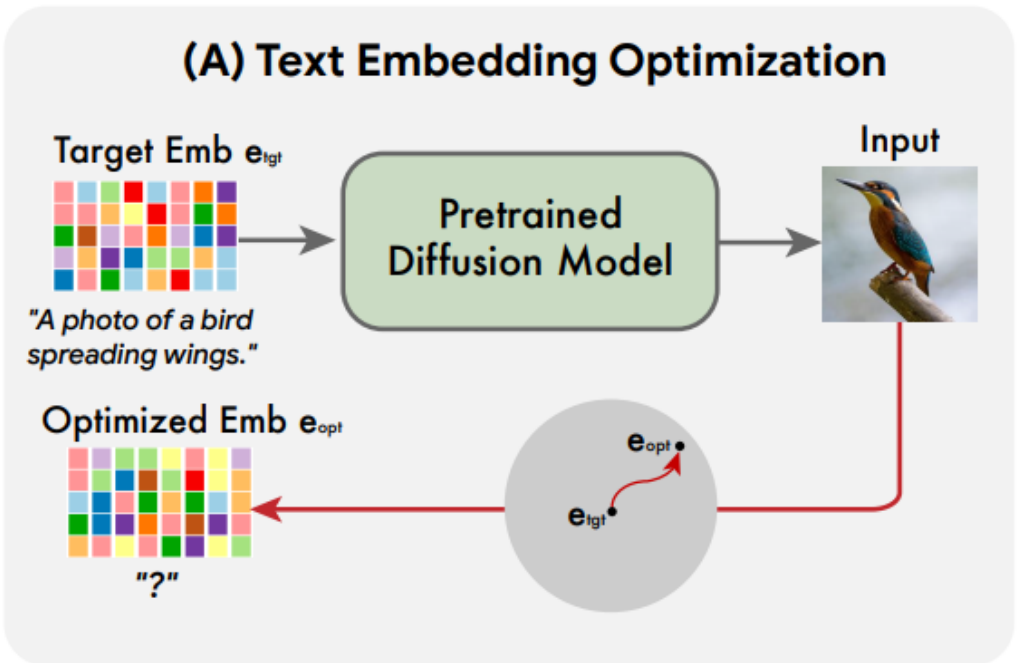
(a) Inputs (b) Real-time rendering (c) Low bitrate streaming



NERFPlayer: Our framework inputs the RGB images captured from a camera array or a single moving camera. After offline optimization, our framework can render a novel view and perform temporal interpolation in real time. Our framework is highly configurable. Adopting TensoRF-CP voxel representation in our framework results in low bitrate streaming of high-quality rendering.

IMAGIC

COMPUTER VISION RESEARCH PAPERS OF THE WEEK 2022



Imagic, a semantic image-altering technique based on Imagen that addresses all the aforementioned issues. Their approach can carry out complex non-rigid changes on actual high-resolution photographs with just an input image to be changed and a single text prompt indicating the target edit. The output images are well-aligned with the target text and maintain the background, composition, and general structure of the source image. Imagic is capable of many alterations, including style adjustments, color changes, and object additions, in addition to more complicated changes. Some examples are shown in the figure below.