Урок 22. Как нейросеть создаёт видео?

9 класс · Урок 22 из 34 · 45 минут

Вся программа 9 класса

Генерация видео нейросетью. Урок для 9 класса о первом кадре, движении и каскадной диффузии. Учимся составлять раскадровку и проверять результат.

Содержание урока

- 01. Цель урока
- 02. Ход урока
- 03. Этап 1. Вступление: Кадр оживает
- 04. Этап 2. Как модель сохраняет согласованность кадров
- 05. Этап 3. Работа над раскадровкой
- 06. Этап 4. Просмотр и монтаж
- 07. Этап 5. Главное и проверка понимания

01. Цель урока

Понять, как нейросеть создаёт последовательность связанных кадров, и научиться задавать движение и проверять его в коротком ролике.

02. Ход урока

Этап	Время
1. Вступление: Кадр оживает	5 мин
2. Как модель сохраняет согласованность кадров	12 мин
3. Работа над раскадровкой	16 мин
4. Просмотр и монтаж	8 мин
5. Главное и проверка понимания	4 мин
Всего	45 мин

03. Этап 1. Вступление: Кадр оживает 5 мин

ГОВОРИМ УЧЕНИКАМ

Для заставки школьной игры нужен кадр, в котором фишка делает шаг с одной клетки на другую. В первом кадре у неё треугольная шапка, а в последнем шапка вдруг круглая. По отдельности обе картинки аккуратные. Почему при просмотре всё равно кажется, что фишку подменили?

Сегодня разберём, как устроена генерация видео нейросетью. После этого попробуем спланировать ролик так, чтобы зритель замечал задуманное движение и узнавал один и тот же предмет в разных кадрах.

Для учителя

Попросите назвать одно изменение, которое задумал автор, и одно случайное. Запишите наблюдения перед переходом к объяснению.

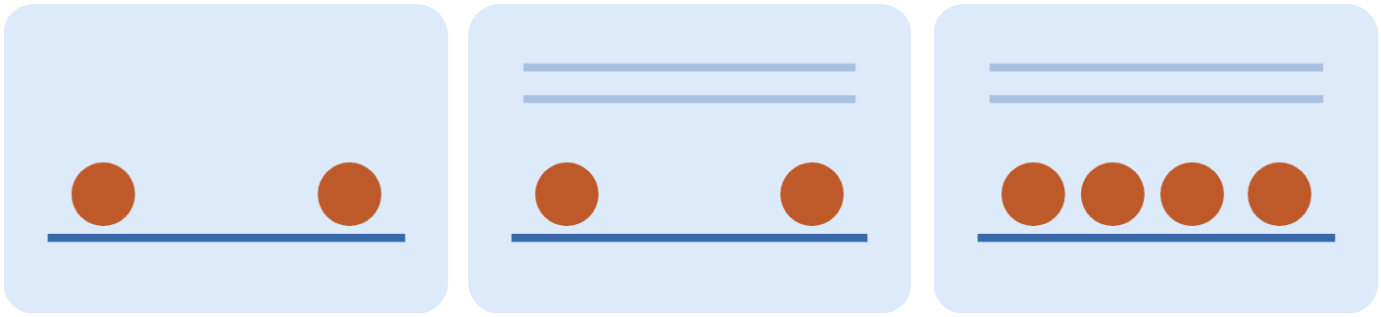
04. Этап 2. Как модель сохраняет согласованность кадров 12 мин

ГОВОРИМ УЧЕНИКАМ

В заставке должна меняться позиция фишки, а форма шапки помогает узнать её. Если шапка меняется сама по себе, зритель может решить, что в кадре уже другой предмет. Поэтому модель должна согласовать кадры между собой. При обучении на видеозаписях она встречает примеры того, как предмет выглядит до движения, во время него и после. Когда модель создаёт ролик, она строит кадры с учётом соседних моментов.

Можно начать с текста: «Камера плавно приближается к витрине, свет гаснет в конце». Тогда модель сама предлагает начальную композицию. Если расположение предметов и их вид уже придуманы, лучше взять первый кадр. Это исходное изображение, на которое модель опирается, когда строит начало ролика. Запрос задаёт, что движется и как. Эффект, например приближение камеры, меняет то, как зритель видит сцену. При движении самого предмета меняется его положение относительно других предметов. Разницу легко заметить, если следить за тем, что находится на заднем плане.

У диффузионных моделей один из способов создать ролик начинается с шума. Модель постепенно убирает шум из кадров, опираясь на запрос. Каскадом называют несколько моделей, которые улучшают результат по очереди. Сначала получается черновой ролик, где уже понятно, что и куда движется. Дальше одни ступени добавляют детали внутри каждого кадра, другие строят промежуточные кадры, чтобы движение выглядело плавнее. Ступени могут чередоваться. Если предмет исчез уже в черновике, повышение чёткости само по себе его не вернёт.



Черновик | Больше деталей | Больше кадров

Точки показывают положение предмета в разные моменты.

Реальный случай. Ролик из нескольких ступеней

В октябре 2022 года исследователи Google представили Imagen Video. Их система сначала строила видео по текстовому запросу. Затем модели по очереди повышали разрешение кадров и добавляли кадры между уже готовыми. Исследователи показали, как можно поручить построение движения и доработку деталей разным ступеням одной системы.

05. Этап 3. Работа над раскадровкой 16 мин

ГОВОРИМ УЧЕНИКАМ

Короткие сгенерированные планы пригодятся, чтобы быстро показать идею для трейлера или проверить, с какого ракурса понятнее выглядит сцена. В длинном ролике труднее следить, чтобы предметы и их детали не менялись сами собой. Поэтому мы начнём с одного движения и проверим, как оно выглядит по ходу сцены.

Задание. Придумайте трёхсекундную заставку

Выберите короткую сцену для события, к которому ваш класс мог бы сделать ролик. Пусть в сцене будет один движущийся предмет и один предмет, который остаётся на месте. Работайте втроём.

1. Набросайте первый, средний и последний кадр. На каждом отметьте, где стоят оба предмета.
2. Напишите запрос для генератора. Решите, понадобится ли ему ваш первый кадр или хватит описания. Укажите движение предмета либо движение камеры и одну деталь, которая должна сохраняться.
3. Закройте средний кадр. Передайте соседям запрос и остальные два кадра. Пусть они предложат свой путь между началом и концом и набросают средний кадр. Сравните две траектории. Если они совпали, придумайте ещё одну, которая

подходит под исходный запрос. Затем уточните фразу так, чтобы одна из траекторий стала предпочтительной.

Результат: раскадровка, уточнённый запрос, две возможные траектории и ограничение, которое помогает выбрать нужную.

Для учителя

Например, если между началом и концом камера должна оставаться неподвижной, положение дальнего предмета останется тем же. Различия между рисунками показывают, где описание можно было понять по-разному. Это опыт с запросом, а работу модели класс изучает по схеме каскада и реальному примеру.

06. Этап 4. Просмотр и монтаж 8 мин

ГОВОРИМ УЧЕНИКАМ

Первый кадр помогает задать вид сцены, но проверить совпадение нужно и в середине, и в конце ролика. Если объект меняет цвет каждый раз, когда проходит за другим предметом, проблема в согласованности сцены. Если он движется верно, но скачками, сначала проверьте плавность. Отдельно проверьте детали: при размытом исходнике более чёткая версия может показать правдоподобный рисунок, которого в исходной сцене не было видно.

Разбор. Какую версию отправить в монтаж

Редактор получил два варианта вращения модели спутника. В первом хорошо видны панели, но одна панель на секунду исчезает. Во втором панели остаются на месте, зато движение немного дёргается. Что бы вы отправили на доработку перед показом на уроке астрономии? Назовите проверку после исправления и один случай, когда ваш выбор мог бы измениться.

Для учителя

Для учебного показа важна неизменность устройства спутника, поэтому удобно взять второй вариант и проверить движение в замедленном просмотре. Для другого замысла, например намеренной покадровой стилизации, рывки могут быть приемлемы. Оценивайте аргументы, связанные с целью ролика.

07. Этап 5. Главное и проверка понимания 4 мин

ГОВОРИМ УЧЕНИКАМ

Когда видео выглядит странно, назовите, что именно нарушилось. Тогда вы сможете решить, стоит ли уточнить сцену в запросе, проверить промежуточные кадры или заняться чёткостью.

Главное за урок

- Первый кадр задаёт исходный вид сцены, а запрос описывает ожидаемое изменение.
- В каскаде разные ступени могут отвечать за движение, частоту кадров и детали.
- Чтобы оценить видео, надо проверить, как меняются предметы во времени.

Итоговые вопросы

1. Как по фону понять, приближается камера или к камере едет предмет?
2. Ролик плавный, но табличка на двери превращается в другую. Что нужно проверить в первую очередь и почему?
3. Движение верное, но кадры мелкие и редкие. Какие две ступени каскада помогут и что изменит каждая?

Для учителя

1. При приближении камеры меняется вид фона, а при движении предмета с неподвижной камерой фон остаётся на месте. 2. Сравнить табличку в нескольких кадрах и вернуться к согласованности сцены. Повышение разрешения лишь сделает видимые отличия резче. 3. Ступень пространственного уточнения добавит детали внутри кадров, а временного — промежуточные кадры.

Источники

- Google Research. Imagen Video · 05.10.2022, материал старше 2024 года. «a sequence of interleaved spatial and temporal video super-resolution models» — в каскаде чередуются модели для кадров и движения.