

عنوان مقاله:

تولید صحنه از روی توصیف متنی با استفاده از یادگیری عمیق

محل انتشار:

دوازدهمین کنفرانس ملی و دومین کنفرانس بین المللی بینایی ماشین و پردازش تصویر ایران (سال: 1400)

تعداد صفحات اصل مقاله: 7

نویسندگان:

پرتو چرخ کار - دانش آموخته ی کارشناسی ارشد مهندسی فناوری اطلاعات، آزمایشگاه پژوهشی یادگیری عمیق، گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشکدگان فارابی، دانشگاه تهران، ایران

کاظم فولادی قلعه - استادیار، گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشکدگان فارابی، دانشگاه تهران، ایران

خلاصه مقاله:

مسئله ی تبدیل توصیف متنی به تصویر از جمله ی مسائلی است که امروزه بسیار مورد توجه قرار گرفته است؛ چرا که داده های تصویری نسبت به داده های متنی برای طیف گسترده تری از افراد از جمله کودکان قابل فهم است و همچنین سرعت انتقال و دریافت مفاهیم از داده های تصویری نسبت به داده های متنی بالاتر است. مدلی که بتواند تبدیل متن به تصویر را با کیفیت قابل قبول انجام دهد می تواند در کاربرد های مختلفی مورد استفاده قرار بگیرد. شبکه های عصبی عمیق با قدرت بازنمایی بالای داده ها مهم ترین ابزار برای انجام این تبدیل هستند. در مدل معرفی شده در این مرجع از شبکه های عصبی مختلفی مانند شبکه ی عصبی کانولوشنال، بازگشتی و بازگشتی کانولوشنال استفاده شده است. مراحل تبدیل توصیف متنی به تصویر شامل (۱) کدگذاری متن و صحنه و (۲) کدگشایی اشیا و ویژگی های آن ها می شود. در این مدل از شبکه های کانولوشنال ResNet-۵۰ و ResNeXt-۵۰ برای کدگذاری صحنه ی تولیدی استفاده شده است و نتایج حاصل از آن ها مقایسه شده است. آموزش و ارزیابی مدل بر روی مجموعه داده ی AbstractScenes انجام شده است. صحنه های تولید شده توسط مدل کیفیت مطلوبی دارند و ارزیابی کمی نتایج نیز عملکرد قابل قبولی را نشان می دهد.

کلمات کلیدی:

تصاویر انتزاعی، توصیف متنی، تولید صحنه، ساخت تصویر

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1428769>

