

عنوان طرح*: تولید تصاویر با روش generative همراه با متادیتا

تعریف مسئله، ضرورت انجام و اهداف طرح*:

مدلهای هوش مصنوعی برای توسعه، نیازمند حجم زیادی از تصاویر و برچسب متناظر برای آموزش و ارزیابی می‌باشند، لذا این حجم از داده با وجود منابع اینترنتی و پایگاه‌های دیتاست، داده تمیز نبوده و یا مطابق نیازمندی نیست در نتیجه لازم است از روشهای نوین همچون GenAI برای جبران این کمبود بهره جست. افزون بر این نه تنها وجود تصاویر بلکه وجود برچسب متناظر (label) با تصویر نیز پراهمیت است.

مشخصات فنی و عملیاتی*:

برای ورودی، رزلوشن تصویر و تعداد تصویر مورد نیاز داده شده و سپس یک جمله توصیفی از تصویر مورد انتظار و صحنه، شامل اشیای داخل صحنه و فضا عنوان شود برای مثال وجود ماشین خاکستری در فضای شهری و آب و هوای ابری سپس چندین تصویر تولید و موقعیت اشیا درخواستی در اینجا ماشین با فرمتهای استاندارد همچون Yolo و Pascal Voc در قالب فایل متنی ذخیره گردد. حالت بعدی تولید ویدیو از سناریو توصیفی و تولید برچسبهای متناظر برای هر فریم از ویدیو میباشد. در حال حاضر مدلهای معروفی برای تولید تصویر و ویدیو در دنیا توسعه داده شده‌اند ولی نیازمندی و چالش اصلی، گرفتن فایل توصیفی این تصاویر شامل گزارش موقعیت اشیا موجود در صحنه می باشد.

خروجی‌های مورد انتظار (دستاوردهای فنی و تولیدات علمی)*:

وجود مدل یا کد با امکان تغییر و بهبود برای تولید تصاویر و فایل توصیفی متناظر با سرعت حداقل پنج فریم بر ثانیه

محدودیتها و قیود:

قابلیت اجرای روی سیستمهای دسکتاپ با پردازندههای گرافیکی مرسوم (مانند RTX 2060, 3070)