# Review 73: [Short] Diffusion Models for Zero-Shot Open-Vocabulary Segmentation

**Paper: https://arxiv.org/abs/2301.05221v2**

המאמר מציע שיטת סגמנטציה zero-shot בעלת מילון פתוח, כלומר אין לנו מכלול סגור קטגוריות אלא כל תמונה מקבלת סגמנטציה עם הקטגוריות משלה. כמובן שהשיטה לא דורשת שום אימון מקדים (zero-shot). זה מתבצע באמצעות מודל Text2Image דיפוזיה (DDPM) גנרטיבי

השיטה המוצעת מכילה שני שלבים:

בשלב הראשון (אימון מקדים) המחברים מגנרטים מספר רב של תמונות לקטגוריה(פרומפט) מסוימת (כמו תמונה טובה של חתול או כלב) בעזרת מודל דיפוזיה מאומן. לאחר מכן הם בונים ״מסד טיפוסים של פיצרים״ של כל קטגוריה שכולל הסגמנטציה של האובייקטים (וגם רקע) באמצעות מודל סגמנטציה מאומן.

בשלב השני מפיקים את הפיצ'רים של תמונה באמצעות CLIP ומשווים אותה עם אב הטיפוסים של הפיצ'רים שישי במסד הנתונים. הקטגוריה והסגמנטציה הדומה ביותר ממסד הפיצ'רים נבחרת.

HuggingFace: https://huggingface.co/papers/2306.09316

Arxiv: https://arxiv.org/pdf/2306.09316.pdf