# Review 118: [Short] Seeing through the Brain: Image Reconstruction of Visual Perception from Human Brain Signals, 08.08.23 https://huggingface.co/papers/2308.02510

**Paper: https://arxiv.org/abs/2208.03666v4**

מכונה שיודעת לקרוא את המחשבות שלנו? האם זה עדיין בגדר החלום או שאנחנו כבר מתקרבים לפתרון? היום ב-shorthebrewpapereviews# סוקרים מאמר שבנה מודל לחיזוי (שחזור) תמונה שמראים לאדם מאות electroencephalogram (EEG) המוקלט מהמוח שלו.   
  
המאמר מאמץ גישה משולבת לעיבוד של אות EEG: מצד אחד מנסים להפיק מהאות פיצ'רים עדינים(fine-grained) של התמונה בדמות מפת בולטות (saliency map) המפיקה את הפיצ'רים הויזואלים החשובים של התמונה (silhouette).   
  
מצד שני מפיקים מהאות גם את הפיצ'רים הגסים של התמונה (ייצוג הכותרת שלה). שני הפיצ'רים אלו מזינים למודל דיפוזיה לטנטי (כמו Stable Diffusion) שמטרתו לשחזר את התמונה. הפיצ'רים העדינים (מפת בולטות) מחושבת בשני שלבים.   
  
בשלב הראשון מחשבים את הייצוג הלטנטי של אות ה-EEG עם למידה ניגודית (מקרבים ייצוגים של אותות EEG לתמונות דומות ומקרבים את אלו לתמונות לא דומות). בשלב השני מאמנים GAN מבוסס על hinge loss (כן עדיין משתמשים בהם) כדי ליצור מפת בולטות של התמונה (הדגימות ה״אמיתיות״ כאן הן התמונות שמראים אותן לאנשים).   
  
הפיצ'רים הגסים מחושבים באופן הבא: יוצרים כותרת של התמונה עם מודל מאומן BLIP (מוקפאת) ומעבירים דרך CLIP כדי ליצור את ייצוגה. ואז מאמנים מודל כך הייצוג הגס המופק מהאות יהיה קרוב לייצוג של כותרת התמונה. ואז מכניסים את מפת הבולטות יחד עם ייצוג הכותרת של התמונה למודל דיפוזיה לטנטי כדי לשחזר את התמונה (האנקודר והדקודר מוקפאים). זה כל הקסם בגדול…