

Evolučné algoritmy pre generovanie kreatívnych obrazov

Yingtao Tian, David Ha

Marek Bavoľár, Matúš Konopinský

Úvod

- **Motivácia:**
 - Optimalizovať geometrické tvary tak, aby reprezentovali konkrétne aj abstraktné pojmy
 - Využitie evolučných stratégií (ES) namiesto gradientných metód
- **Cieľ práce:** Ukázať, že moderné evolučné stratégie (Evolution Strategies, ES) môžu efektívne vytvárať obrazy reprezentujúce konkrétne objekty aj abstraktné myšlienky

Moderné evolučné stratégie (ES)

- **ES** je konkrétny typ **EA**, špecializovaný na optimalizáciu reálnych parametrov a adaptívne riadenie mutácií
- **Princíp:** Mierne upravovať (mutovať) parametre riešenia a sledovať, ktoré zmeny vedú k lepšiemu výsledku
- **Využité technológie:**
 - **PGPE** (Policy Gradient with Parameter-based Exploration): Efektívne učenie sa pomocou riadeného šumu
 - **ClipUp** optimizer: Jednoduchý optimalizátor, ktorý orezáva veľkosť zmien parametrov, aby sa stabilizovalo učenie
- **Cieľ:** Optimalizovať obrazy z trojuholníkov tak, aby zodpovedali textovým popisom podľa CLIP modelu

Reprezentácia jedinca

- **Jedinec** = súbor trojuholníkov, ktoré sa kreslia na plátno
- **Každý trojuholník má:**
 - Tri vrcholy (x_1, y_1) , (x_2, y_2) , (x_3, y_3)
 - Farbu (R, G, B) a priehľadnosť (A)
- **Vnútná reprezentácia riešenia:** všetky tieto parametre pre všetky trojuholníky
- Trojuholníky sa vrstvia, čím vytvárajú komplexné obrazy

Cieľ evolúcie a fitness funkcia

- **Cieľ:** Nájsť obrazy, ktoré najlepšie reprezentujú daný textový popis
- **Fitness funkcia:** Hodnotí podobnosť medzi **generovaným obrazom** a **textovým promptom** pomocou modelu **CLIP**
- **CLIP** je model od OpenAI, ktorý bol trénovaný na obrovskom množstve dvojíc text - obrázok (internetové obrázky + ich popisy)
- **Úlohou CLIPu je:**
 - Premeniť textové vstupy aj obrázky na vektory v rovnakom priestore (embedding space)
 - V tomto priestore môžeme merať, ako veľmi "zodpovedá" obrázok textovému popisu

Evolučné operátory

- **Mutácia:**

- Malé náhodné zmeny v parametroch trojuholníkov (pozícia, farba, veľkosť)
- Pridanie alebo odstránenie trojuholníka

- **Crossover:**

- Kombinovanie trojuholníkov od viacerých rodičov
- Mierne využívaný kvôli zachovaniu stability riešenia

- **Repair kroky:**

- Odstránenie malých alebo neefektívnych tvarov (veľmi malých alebo nevýrazných trojuholníkov)
- Jemné doladenie parametrov po mutáciách a crossoveri

Adaptívne stratégie v optimalizácii

- **Adaptívne mutácie:**
 - Väčšie zmeny pri stagnácii
 - Menšie zmeny pri zlepšení
- **ClipUp:**
 - Udržiava stabilnú veľkosť zmien bez neprimeraného skoku
- **Výsledok:**
 - Efektívne a stabilné učenie aj v prostrediach s veľkým šumom

Tvorba konkrétnych obrazov

- **Príklady:** "motýl", "auto", "strom"
- Evolúcia doladuje trojuholníky tak, aby boli obrazy rozpoznateľné a esteticky zaujímavé
- CLIP pomáha hodnotiť, či je výsledok správne interpretovaný

Target Image	10 Triangles	25 Triangles	50 Triangles	200 Triangles
				
"Darwin"	Fitness = 96.82%	Fitness = 99.25%	Fitness = 99.51%	Fitness = 99.75%
				
"Mona Lisa"	Fitness = 98.02%	Fitness = 99.30%	Fitness = 99.62%	Fitness = 99.80%
				
"Anime Face"	Fitness = 94.97%	Fitness = 98.17%	Fitness = 98.80%	Fitness = 99.07%
				
"Velociraptor"	Fitness = 95.76%	Fitness = 98.03%	Fitness = 99.00%	Fitness = 99.25%
				
"Landscape"	Fitness = 97.07%	Fitness = 98.83%	Fitness = 99.08%	Fitness = 99.25%
				
"Impressionism"	Fitness = 98.82%	Fitness = 99.23%	Fitness = 99.34%	Fitness = 99.48%

Tvorba abstraktných konceptov

- **Príklady:** "sila", "sloboda", "smútok"
- Evolúcia sa snaží zachytiť emocionálny alebo symbolický význam
- Výsledné obrazy bývajú viac umelecké a otvorené interpretácii

Prompt	Evolved Results	Evolution Process
“Self”		
“Human”		
“Walt Disney World”		
“The corporate headquarters ...” (full sentence in caption)		

Výhody a výzvy evolučných stratégií

- **Výhody:**

- Nezávislosť od gradientu
- Odolnosť voči šumu
- Podpora diverzity riešení

- **Výzvy:**

- Vyššia výpočtová náročnosť
- Pomalšia konvergencia
- Citlivé nastavenie parametrov

Záver

- **Zistenia:** Evolučné stratégie dokážu generovať esteticky aj konceptuálne kvalitné obrazy
- **Význam:** Spájanie textu a obrazu pomocou CLIP a ES otvára nové možnosti kreatívneho generovania obsahu
- **Budúcnosť:** Podobné prístupy môžu nájsť využitie v umení, dizajne aj automatickej tvorbe obsahu

Článok: [Modern Evolution Strategies for Creativity: Fitting Concrete Images and Abstract Concepts](#)