Differentiable Quality Diversity

Anonymous Author(s)
Affiliation
Address
email

Abstract

Quality diversity (QD) is a growing branch of stochastic optimization research that studies the problem of generating an archive of solutions that maximize a given objective function but are also diverse with respect to a set of specified measure functions. However, even when these functions are differentiable, QD algorithms treat them as "black boxes", ignoring gradient information. We present the differentiable quality diversity (DQD) problem, a special case of QD, where both the objective and measure functions are first order differentiable. We then present MAP-Elites via Gradient Arborescence (MEGA), a DQD algorithm that leverages gradient information to efficiently explore the joint range of the objective and measure functions. Results in two QD benchmark domains and in searching the latent space of a StyleGAN show that MEGA significantly outperforms state-of-the-art QD algorithms, highlighting DQD's promise for efficient quality diversity optimization when gradient information is available.

1 Introduction

We introduce the problem of differentiable quality diversity (DQD) and propose the MAP-Elites via Gradient Arborescence (MEGA) algorithm as the first DQD algorithm.

Unlike single-objective optimization, quality diversity (QD) is the problem of finding a range of high quality solutions that are diverse with respect to prespecified metrics. For example, consider the problem of generating realistic images that match as closely as possible a target text prompt "Elon Musk", but vary with respect to hair and eye color. We can formulate the problem of searching the latent space of a generative adversarial network (GAN) [24] as a QD problem of discovering latent codes that generate images maximizing a matching score for the prompt "Elon Musk", while achieving a diverse range of measures of hair and eye color, assessed by visual classification models [47]. More generally, the QD objective is to maximize an objective f for each output combination of measure functions m_i.

While our example problem can be formulated as a QD problem, all current QD algorithms treat the objective f and measure functions m_i as a black box. However, current QD algorithms fail to take advantage of the fact that both f and m_i are end-to-end differentiable neural networks. Our proposed differentiable quality diversity (DQD) algorithms leverage first-order derivative information to significantly improve the computational efficiency of solving a variety of QD problems where f and m_i are differentiable.

To solve DQD, we introduce the concept of a *gradient arborescence*. Like gradient ascent, a gradient arborescence makes greedy ascending steps based on the objective f. Unlike gradient ascent, a gradient arborescence encourages exploration by branching via the measures m_i.

Our work makes four main contributions. 1) We introduce and formalize the problem of differentiable quality diversity (DQD). 2) We propose two DQD algorithms: Objective and Measure Gradient MAP-Elites via Gradient Arborescence (OMG-MEGA), an algorithm based on MAP-Elites [13],

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

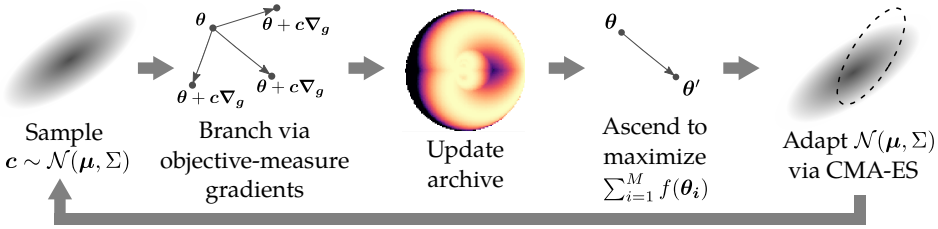

Figure 1: An overview of the Covariance Matrix Adaptation MAP-Elites via Gradient Arborescence (CMA-MEGA) algorithm. The algorithm leverages a gradient arborescence to branch in objective-measure space, while dynamically adapting the gradient steps to maximize archive improvement.

which branches based on the measures m_i but ascends based on the objective function f; and Covariance Matrix Adaptation MEGA (CMA-MEGA) which is based on the CMA-ME [18] algorithm, and which branches based on the objective-measure space but ascends based on maximizing the QD objective (Fig. 1). Both algorithms search directly in measure space and leverage the gradients of f and m_i to form efficient parameter space steps in $\boldsymbol{\theta}$. 3) We show in three different QD domains (the linear projection, the arm repertoire, and the latent space illumination (LSI) domains), that DQD algorithms significantly outperform state-of-the-art QD algorithms that treat the objective and measure functions as a black box. 4) We demonstrate how searching the latent space of a StyleGAN [33] in the LSI domain with CMA-MEGA results in a diverse range of high-quality images.

2 Problem Definition

Quality Diversity. The quality diversity (QD) problem assumes an objective $f : \mathbb{R}^n \to \mathbb{R}$ in an n-dimensional continuous space $\mathbb{R}^n$ and k measures $m_i : \mathbb{R}^n \to \mathbb{R}$ or, as a joint measure, $\boldsymbol{m} : \mathbb{R}^n \to \mathbb{R}^k$. Let $S = \boldsymbol{m}(\mathbb{R}^n)$ be the measure space formed by the range of m. For each $\boldsymbol{s} \in S$ the QD objective is to find a solution $\boldsymbol{\theta} \in \mathbb{R}^n$ such that $\boldsymbol{m}(\boldsymbol{\theta}) = \boldsymbol{s}$ and $f(\boldsymbol{\theta})$ is maximized.

However, we note that $\mathbb{R}^k$ is continuous, and an algorithm solving the quality diversity problem would require infinite memory to store all solutions. Thus, QD algorithms in the MAP-Elites [40, 13] family approximate the problem by discretizing S via a tessellation method. Let T be the tessellation of S into M cells. We relax the QD objective to find a set of solutions $\boldsymbol{\theta_i}, i \in \{1, \ldots, M\}$, such that each $\boldsymbol{\theta_i}$ occupies one unique cell in T. The occupants $\boldsymbol{\theta_i}$ of all M cells form an archive of solutions. Each solution $\boldsymbol{\theta_i}$ has a position in the archive $\boldsymbol{m}(\boldsymbol{\theta_i})$, corresponding to one out of M cells, and an objective value $f(\boldsymbol{\theta_i})$.

The objective of QD can be rewritten as follows, where the goal is to maximize the objective value for each cell in the archive:

$$\max \sum_{i=1}^{M} f(\boldsymbol{\theta_i}) \tag{1}$$

Differentiable Quality Diversity. We define the differentiable quality diversity (DQD) problem, as a QD problem where both the objective f and measures m_i are *first-order differentiable*.

3 Preliminaries

We present several state-of-the-art derivative-free QD algorithms. Our proposed DQD algorithm MEGA builds upon ideas from these works, while introducing measure and objective gradients into the optimization process.

MAP-Elites and MAP-Elites (line). MAP-Elites [13, 40] first tessellates the measure space S into evenly-spaced grid cells. The upper and lower bounds for $\boldsymbol{m}$ are given as input to constrain S to a finite region. MAP-Elites first samples solutions from a fixed distribution $\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0}, I)$, and populates an initial archive after computing $f(\boldsymbol{\theta})$ and $\boldsymbol{m}(\boldsymbol{\theta})$. Each iteration of MAP-Elites selects λ cells uniformly at random from the archive and perturbs each occupant $\boldsymbol{\theta_i}$ with fixed-variance σ

isotropic Gaussian noise: $\boldsymbol{\theta}' = \boldsymbol{\theta_i} + \sigma\mathcal{N}(\mathbf{0}, I)$. Each new candidate solution $\boldsymbol{\theta'}$ is then evaluated and added to the archive if $\boldsymbol{\theta'}$ discovers a new cell or improves an existing cell. The algorithm continues to generate solutions for a specified number of iterations.

Later work introduced the Iso+LineDD operator [56]. The Iso+LineDD operator samples two archive solutions $\boldsymbol{\theta_i}$ and $\boldsymbol{\theta_j}$, then blends a Gaussian perturbation with a noisy interpolation given hyperparameters σ_1 and σ_2: $\boldsymbol{\theta'} = \boldsymbol{\theta_i} + \sigma_1\mathcal{N}(\mathbf{0}, I) + \sigma_2\mathcal{N}(\mathbf{0}, 1)(\boldsymbol{\theta_i} - \boldsymbol{\theta_j})$. In our paper we refer to MAP-Elites with an Iso+LineDD operator as MAP-Elites (line).

CMA-ME. Covariance Matrix Adaptation MAP-Elites (CMA-ME) [18] combines the archiving mechanisms of MAP-Elites with the adaptation mechanisms of CMA-ES [28]. While MAP-Elites creates new solutions by perturbing existing solutions with fixed-variance Gaussian noise, CMA-ME maintains a full-rank Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ in parameter space $\mathbb{R}^n$. Each iteration of CMA-ME samples λ candidate solutions $\boldsymbol{\theta_i} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, evaluates each solution, and updates the archive based on the following rule: if there is a previous occupant $\boldsymbol{\theta_p}$ at the same cell, we compute $\Delta_i = f(\boldsymbol{\theta_i}) - f(\boldsymbol{\theta_p})$, otherwise if the cell is empty we compute $\Delta_i = f(\boldsymbol{\theta_i})$. We then rank the sampled solutions by increasing improvement Δ_i, with an extra criteria that candidates discovering new cells are ranked higher than candidates that improve existing cells. We then update $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ with the standard CMA-ES update rules based on the improvement ranking. CMA-ME restarts when all λ solutions fail to change the archive. On a restart we reset the Gaussian $\mathcal{N}(\boldsymbol{\theta_i}, I)$, where $\boldsymbol{\theta_i}$ is an archive solution chosen uniformly at random, and all internal CMA-ES parameters. In the supplemental material, we derive, for the first time, a natural gradient interpretation of the CMA-ME's improvement ranking mechanism, based on previous theoretical work on CMA-ES [2].

4 Algorithms

We present two variants of our proposed MEGA algorithm: OMG-MEGA and CMA-MEGA. We form each variant by adapting the concept of a gradient arborescence to the MAP-Elites and CMA-ME algorithms, respectively. Finally, we introduce a third algorithm, OG-MAP-Elites, which operates only on the gradients of the objective, as a baseline.

OMG-MEGA. We first derive the Objective and Measure Gradient Map-Elites via Gradient Arborescence (OMG-MEGA) algorithm from MAP-Elites.

First, we observe how gradient information could benefit a QD algorithm. Note that the QD objective is to explore the measure space, while maximizing the objective function f. We observe that maximizing a linear combination of measures : $\sum_{j=1}^{k} c_j m_j(\boldsymbol{\theta})$, where $\boldsymbol{c}$ is a k-dimensional vector of coefficients, enables movement in a k-dimensional measure space. Adding the objective function f to the linear sum enables movement in an objective-measure space. Maximizing g with a positive coefficient of f results in steps that increasing f, while the direction of movement in the measure space is determined by the sign and magnitude of the coefficients c_j.

$$g(\boldsymbol{\theta}) = |c_0| f(\boldsymbol{\theta}) + \sum_{j=1}^{k} c_j m_j(\boldsymbol{\theta}) \tag{2}$$

We can then derive a direction function that perturbs a given solution $\boldsymbol{\theta}$ based on the gradient: $\boldsymbol{\nabla} g(\boldsymbol{\theta}) = |c_0| \boldsymbol{\nabla} f(\boldsymbol{\theta}) + \sum_{j=1}^{k} c_j \boldsymbol{\nabla} m_j(\boldsymbol{\theta})$. We incorporate the direction function $\boldsymbol{\nabla} g$ to derive a gradient-based MAP-Elites variation operator.

We observe that MAP-Elites samples a cell $\boldsymbol{\theta_i}$ and perturbs the occupant with Gaussian noise: $\boldsymbol{\theta'} = \boldsymbol{\theta_i} + \sigma\mathcal{N}(\mathbf{0}, I)$. Instead, we sample coefficents $\boldsymbol{c} \sim \mathcal{N}(\mathbf{0}, \sigma_g I)$ and update:

$$\boldsymbol{\theta'} = \boldsymbol{\theta_i} + \mid c_0 \mid \boldsymbol{\nabla} f(\boldsymbol{\theta_i}) + \sum_{j=1}^{k} c_j \boldsymbol{\nabla} m_j(\boldsymbol{\theta_i}) \tag{3}$$

The value σ_g acts as a learning rate for the gradient step. To balance the contribution of each function, we normalize each gradient. Other than our new gradient-based operator, OMG-MEGA is identical to MAP-Elites.

CMA-MEGA. Next, we derive the Covariance Matrix Adaptation MAP-Elites via Gradient Arborescence algorithm from CMA-ME. Fig. 1 shows an overview of the algorithm.

First, we note that we sample $\boldsymbol{c}$ in OMG-MEGA from a fixed-variance Gaussian. However, it would be beneficial to select $\boldsymbol{c}$ based on how $\boldsymbol{c}$, and the subsequent gradient step on $\boldsymbol{\theta}$, improve the QD objective defined in equation 1.

We frame the selection of $\boldsymbol{c}$ as an optimization problem with the objective of maximizing archive improvement. We model a distribution of coefficients $\boldsymbol{c}$ as a $k+1$-dimensional Gaussian $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$. Given a $\boldsymbol{\theta}$, we can sample $\boldsymbol{c} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, compute $\boldsymbol{\theta'}$ via Eq. 3, and adapt $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ towards the direction of maximum increase of the QD objective.

We follow an evolution strategy approach to model and dynamically adapt the sampling distribution of coefficients $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$. We sample a population of λ coefficients from $\boldsymbol{c_i} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ and generate λ solutions $\boldsymbol{\theta_i}$. We then compute Δ_i from CMA-ME's improvement ranking for each candidate solution $\boldsymbol{\theta_i}$. By updating $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ with CMA-ES update rules for the ranking Δ_i, we dynamically adapt the distribution of coefficients $\boldsymbol{c}$ to maximize increase of the QD objective.

Algorithm 1 shows the pseudocode for CMA-MEGA. In line 3 we evaluate the current solution and compute an objective value f, a vector of measure values $\boldsymbol{m}$, and gradient values. As we dynamically adapt the coefficients $\boldsymbol{c}$, we normalize the objective and measure gradients (line 4) for stability. Because the measure space is tessellated, the measures $\boldsymbol{m}$ place solution $\boldsymbol{\theta}$ into one of the M cells in the archive. We then add the solution to the archive (line 5), if the solution discovers an empty cell in the archive, or if it improves an existing cell, identically to MAP-Elites.

We then use the gradient information to compute a step that maximizes improvement of the archive. In lines 6-12, we sample a population of λ coefficients from a multi-variate Gaussian retained by CMA-ES, and take a gradient step for each sample. We evaluate each sampled solution $\boldsymbol{\theta'_i}$, and compute the improvement Δ_i (line 11). As in CMA-ME, we specify Δ_i as the difference in the objective value between the sampled solution $\boldsymbol{\theta}_i$ and the existing solution, if one exists, or as the absolute objective value of the sampled solution if $\boldsymbol{\theta_i}$ belongs to an empty cell.

In line 13, we rank the sampled gradients $\boldsymbol{\nabla}_i$ based on their respective improvements. As in CMA-ME, we prioritize exploration of the archive by ranking first by their objective values all samples that discover new cells, and subsequently all samples that improve existing cells by their difference in improvement. We then use the ranking to compute the ascending gradient step as a linear combination of the gradients (line 14), following the recombination weights from CMA-ES [28].

In line 16, CMA-ES adapts the multi-variate Gaussian $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$, as well as internal search parameters $\boldsymbol{p}$, from the improvement ranking of the coefficients. In the supplemental material, we provide a natural gradient interpretation of the improvement ranking rules of CMA-MEGA, where we show that the coefficient distribution of CMA-MEGA approximates natural gradient steps of maximizing a modified QD objective.

CMA-MEGA (Adam). We add an Adam-based variant of CMA-MEGA, where we replace line 15 with an Adam gradient optimization step [35].

OG-MAP-Elites. To show the importance of taking gradient steps in the measure space, as opposed to only taking gradient steps in objective space and directly perturbing the parameters, we derive a variant of MAP-Elites as a baseline based off the recently proposed Policy Gradient Assisted Map-Elites (PGA-ME) algorithm [42]. PGA-ME combines a Gaussian variation operator with a policy gradient operator *only on the objective*. Note that we cannot compare directly against PGA-ME as the algorithm specializes for reinforcement learning. Instead, we simplify PGA-ME to optimize a vanilla objective gradient called Objective-Gradient MAP-Elites (OG-MAP-Elites). Each iteration of OG-MAP-Elites samples λ solutions $\boldsymbol{\theta_i}$ from the archive. Each $\boldsymbol{\theta_i}$ is perturbed with Gaussian noise to form a new candidate solution $\boldsymbol{\theta'_i} = \boldsymbol{\theta_i} + \sigma\mathcal{N}(\mathbf{0}, I)$. OG-MAP-Elites evaluates the solution and updates the archive, exactly as in MAP-Elites. However, OG-MAP-Elites takes one additional step: for each $\boldsymbol{\theta'_i}$, the algorithm computes $\boldsymbol{\nabla} f(\boldsymbol{\theta'_i})$, forms a new solution $\boldsymbol{\theta''_i} = \boldsymbol{\theta'_i} + \eta\boldsymbol{\nabla} f(\boldsymbol{\theta'_i})$ with an objective gradient step, and evaluates $\boldsymbol{\theta''_i}$. Finally, we update the archive with all solutions $\boldsymbol{\theta'_i}$ and $\boldsymbol{\theta''_i}$.

Algorithm 1 Covariance Matrix Adaptation MAP-Elites via Gradient Aborescence (CMA-MEGA)

CMA-MEGA $(evaluate, \boldsymbol{\theta_0}, N, \lambda, \eta, \sigma_g)$

input : An evaluation function $evaluate$ which computes the objective, the measures, gradients of the objective and measures, an initial solution $\boldsymbol{\theta_0}$, a desired number of iterations N, a branching population size λ, a learning rate η, and an initial step size for CMA-ES σ_g.

result : Generate $N\lambda$ solutions storing elites in an archive A.

Initialize solution parameters $\boldsymbol{\theta}$ to $\boldsymbol{\theta_0}$, CMA-ES parameters $\boldsymbol{\mu} = \mathbf{0}$, $\Sigma = \sigma_g I$, and $\boldsymbol{p}$, where we let $\boldsymbol{p}$ be the CMA-ES internal parameters.

for $iter \leftarrow 1$ **to** N **do**

  $f, \boldsymbol{\nabla}_f, \boldsymbol{m}, \boldsymbol{\nabla}_{\boldsymbol{m}} \leftarrow$ evaluate$(\boldsymbol{\theta})$

  $\boldsymbol{\nabla}_f \leftarrow$ normalize$(\boldsymbol{\nabla}_f), \boldsymbol{\nabla}_m \leftarrow$ normalize$(\boldsymbol{\nabla}_m)$

  update_archive$(\boldsymbol{\theta}, f, \boldsymbol{m})$

  for $i \leftarrow 1$ **to** λ **do**

    $\boldsymbol{c} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$

    $\boldsymbol{\nabla}_i \leftarrow c_0 \boldsymbol{\nabla}_f + \sum_{j=1}^{k} c_j \boldsymbol{\nabla}_{m_j}$

    $\boldsymbol{\theta'_i} \leftarrow \boldsymbol{\theta} + \boldsymbol{\nabla}_i$

    $f', *, \boldsymbol{m'}, * \leftarrow$ evaluate$(\boldsymbol{\theta'_i})$

    $\Delta_i \leftarrow$ update_archive$(\boldsymbol{\theta'_i}, f', \boldsymbol{m'})$

  end

  rank $\boldsymbol{\nabla}_i$ by Δ_i

  $\boldsymbol{\nabla}_{\text{step}} \leftarrow \sum_{i=1}^{\lambda} w_i \boldsymbol{\nabla}_{\text{rank[i]}}$

  $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \eta \boldsymbol{\nabla}_{\text{step}}$

  Adapt CMA-ES parameters $\boldsymbol{\mu}, \Sigma, \boldsymbol{p}$ based on improvement ranking Δ_i

  if *there is no change in the archive* **then**

    Restart CMA-ES with $\boldsymbol{\mu} = 0, \Sigma = \sigma_g I$.

    Set $\boldsymbol{\theta}$ to a randomly selected existing cell $\boldsymbol{\theta_i}$ from the archive

  end

end

5 Domains

We select benchmark domains from previous work in the QD literature, but we focus on domains with differentiable objective and measures, where DQD is applicable.

Linear Projection. To show the importance of adaptation mechanisms in QD, the CMA-ME paper [18] introduced a simple domain, where reaching the extremes of the measures is challenging for non-adaptive QD algorithms. The domain forms each measure m_i by a linear projection from $\mathbb{R}^n$ to $\mathbb{R}$, while bounding the contribution of each component $\boldsymbol{\theta}_i$ to the range $[-5.12, 5.12]$.

We note that uniformly sampling from a hypercube in $\mathbb{R}^n$ results in a narrow distribution of the linear projection in $\mathbb{R}$ [18, 31]. Increasing the number of parameters n makes the problem of covering the measure space more challenging, because to reach an extremum $m_i(\boldsymbol{\theta}) = \pm 5.12n$, all components must equal the extremum: $\boldsymbol{\theta}[\boldsymbol{i}] = \pm 5.12$.

We select this domain as a benchmark to highlight the need for adaptive gradient coefficients for CMA-MEGA as opposed to constant coefficients for OMG-MEGA, because reaching the edges of the measure space requires dynamically shrinking the gradient steps.

As a QD domain, the domain must provide an objective. The CMA-ME study [18] introduces two variants of the linear projection domain with an objective based on the sphere and Rastrigin functions from the continuous black-box optimization set of benchmarks [27, 29]. We optimize an $n = 1000$ unbounded parameter space $\mathbb{R}^n$. We provide more detail in the supplemental material.

Arm Repertoire. We select the robotic arm repertoire domain from previous work [13, 56]. The goal in this domain is to find an inverse kinematics (IK) solution for each reachable position of the end-effector of a planar robotic arm with revolute joints. The objective f of each solution is to minimize the variance of the joint angles, while the measure functions are the positions of the end effector in the x and y-axis, computed with the forward kinematics of the planar arm [41]. We selected a 1000-DOF robotic arm.

	LP (sphere)		LP (Rastrigin)		Arm Repertoire		LSI	
Algorithm	QD-score	Coverage	QD-score	Coverage	QD-score	Coverage	QD-score	Coverage
MAP-Elites	1.04	1.17%	1.18	1.72%	1.97	8.06%	13.88	23.15%
MAP-Elites (line)	12.21	14.32%	8.12	11.79%	33.51	35.79%	16.54	25.73%
CMA-ME	1.08	1.21%	1.21	1.76%	55.98	56.95%	18.96	26.18%
OG-MAP-Elites	1.52	1.67%	1.26	1.67%	57.17	58.08%	N/A	N/A
OMG-MEGA	71.58	92.09%	55.90	77.00%	44.12	44.13%	N/A	N/A
CMA-MEGA	75.29	**100.00%**	62.54	**100.00%**	**74.18**	**74.18%**	5.36	8.61%
CMA-MEGA (Adam)	**75.30**	**100.00%**	**62.58**	**100.00%**	73.82	73.82%	**21.82**	**30.73%**

Table 1: Results: Mean QD-score and coverage values after 10,000 iterations for each algorithm per domain.

Latent Space Illumination. Previous work [19] introduced the problem of exploring the latent space of a generative model directly with a QD algorithm. The authors named the problem latent space illumination (LSI). As the original LSI work evaluated non-differentiable objectives and measures, we create a new benchmark for the differentiable LSI problem by generating images with StyleGAN [33] and leveraging CLIP [47] to create differentiable objective and measure functions. We adopt the StyleGAN+CLIP [1] pipeline, where StyleGAN-generated images are passed to CLIP, which in turn evaluates how well the generated image matches a given text prompt. We form the prompts "Elon Musk with short hair." as the objective and for the measures we form the prompts "A person with red hair." and "A man with blue eyes.". The goal of DQD becomes generating faces similar to Elon Musk with short hair, but varying with respect to hair and eye color.

6 Experiments

We conduct experiments to assess the performance of the MEGA variants. In addition to our baseline OG-MAP-Elites, which we propose in section 4, we compare the MEGA variants with the state-of-the-art QD algorithms presented in section 3.

6.1 Experiment Design

Independent Variables. We follow a between-groups design, where the independent variables are the algorithm and the domain (linear projection, arm repertoire, and LSI). We did not run OMG-MEGA and OG-MAP-Elites in the LSI domain; while CMA-MEGA computes the f and m_i gradients once per iteration (line 3 in Algorithm 1), OMG-MEGA and OG-MAP-Elites compute the f and m_i gradients for every sampled solution, making their execution cost-prohibitive for the LSI domain.

Dependent Variables. We measure both the diversity and the quality of the solutions returned by each algorithm. These are combined by the QD-score metric [45], which is defined as the sum of f values of all cells in the archive (Eq. 1). To make the QD-score invariant with respect to the resolution of the archive, we normalize QD-score by the archive size. As an additional metric of diversity we compute the coverage as the number of occupied cells in the archive divided by the total number of cells. We run each algorithm for 20 trials in the linear projection and arm repertoire domains, and for 5 trials in the LSI domain, resulting in a total of 445 trials.

6.2 Analysis

Table 1 shows the metrics of all the algorithms, averaged over 20 trials for the benchmark domains and over 5 trials for the LSI domain. We conducted a two-way ANOVA to examine the effect of algorithm and domain (linear projection (sphere), linear projection (Rastrigin), arm repertoire) on the QD-Score. There was a statistically significant interaction between the search algorithm and the domain ($F(12, 399) = 6453.43, p < 0.001$). Simple main effects analysis with Bonferroni corrections showed that CMA-MEGA and OMG-MEGA performed significantly better than each of the baselines in the sphere and Rastrigin domains, with CMA-MEGA significantly outperforming OMG-MEGA. CMA-MEGA also outperformed all the other algorithms in the arm repertoire domain.

We additionally conducted a one-way ANOVA to examine the effect of algorithm on the LSI domain. There was a statistically significant difference between groups ($F(4, 20) = 260.64, p < 0.001$). Post-hoc pairwise comparisons with Bonferroni corrections showed that CMA-MEGA (Adam) significantly

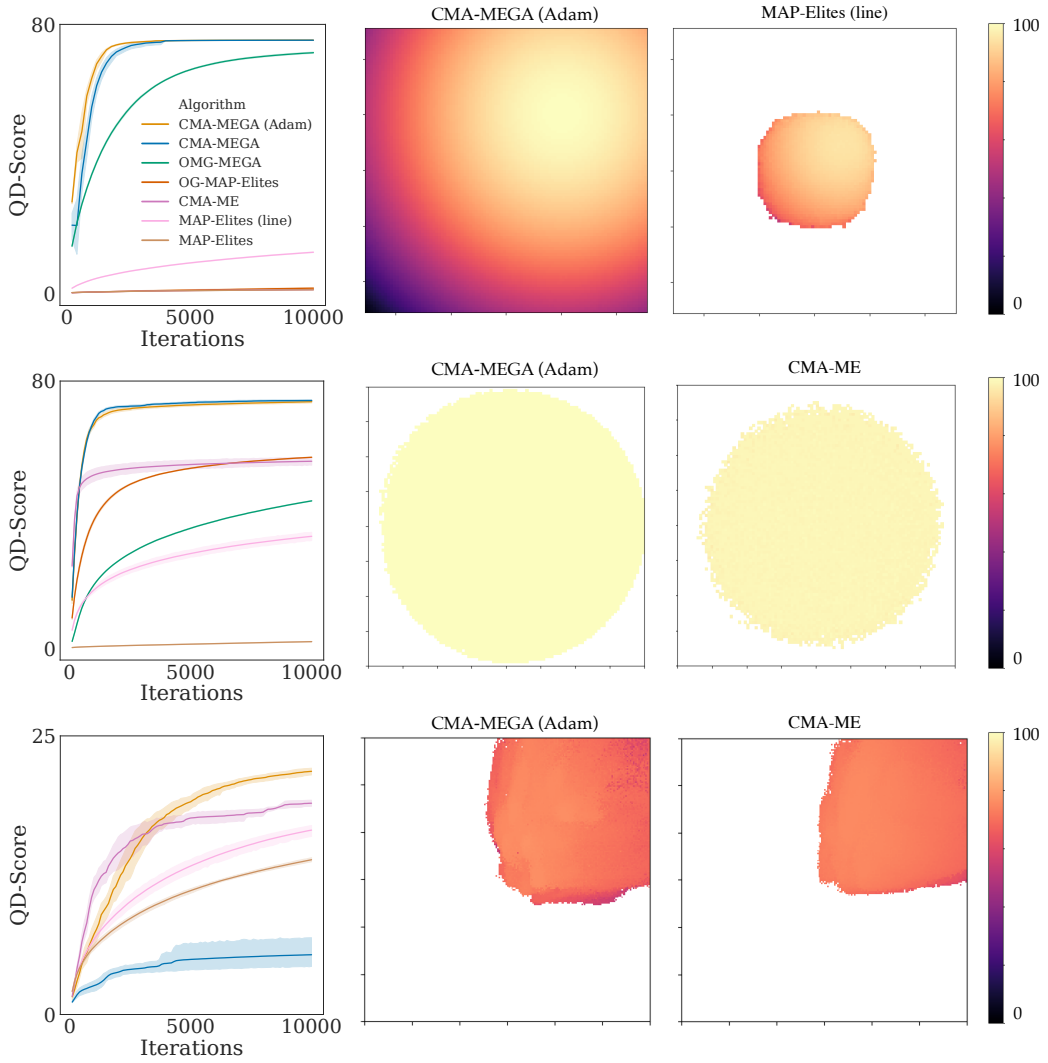

Figure 2: QD-Score plot with 95% confidence intervals and heatmaps of generated archives by CMA-MEGA (Adam) and the strongest derivative-free competitor for the linear projection sphere (top), arm repertoire (middle), and latent space illumination (bottom) domains.

outperformed all other algorithms, while CMA-MEGA without the Adam implementation had the worst performance.

Both OMG-MEGA and CMA-MEGA variants perform well in the linear projection domain, where the objective and measure functions are additively separable, and the partial derivatives with respect to each parameter independently capture the steepest change of each function. We observe that OG-MAP-Elites performs poorly in this domain. Analysis shows that the algorithm finds a nearly perfect best solution for the sphere objective, but it interleaves following the gradient of the objective with exploring the archive as in standard MAP-Elites, resulting in smaller coverage of the archive.

In the arm domain, OMG-MEGA manages to reach the extremes of the measure space, but the algorithm fails to fill in nearby cells. OG-MAP-Elites performs significantly better than OMG-MEGA, because the top-performing solutions in this domain tend to be concentrated in an "elite hypervolume" [56]; moving towards the gradient of the objective finds top-performing cells, while applying isotropic perturbations to these cells fills in nearby regions in the archive. CMA-MEGA variants retain the best performance in this domain. Fig. 1 shows a high-precision view of the CMA-MEGA (Adam) archive for the arm repertoire domain.

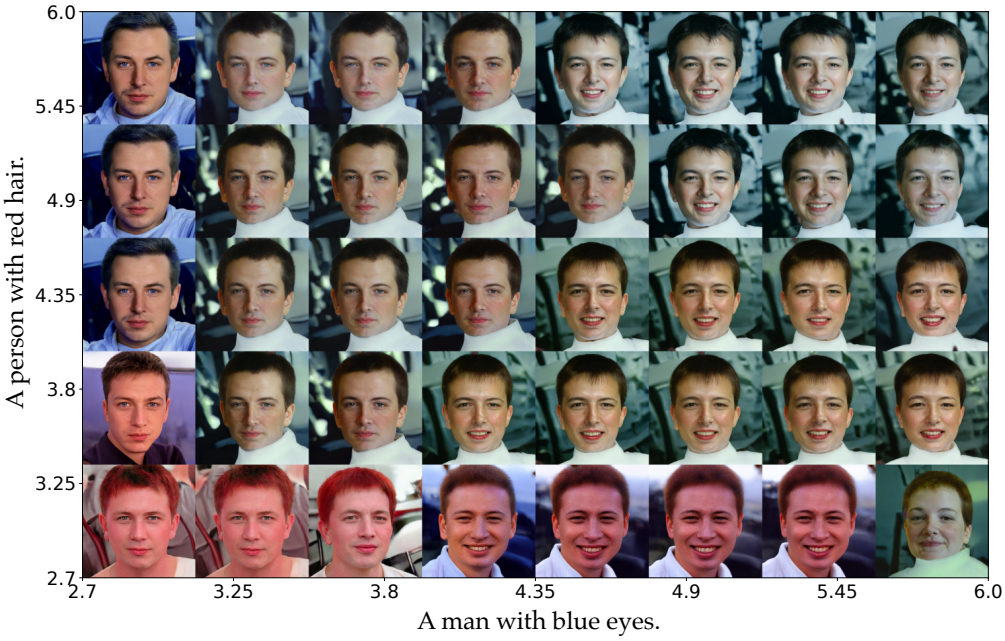

Figure 3: Result of latent space illumination for objective "Elon Musk with short hair." and for measures "A person with red hair." and "A man with blue eyes.". The axes values indicate the score returned by the CLIP model, where lower score indicates a better match.

We did not observe a large difference between the CMA-MEGA (Adam) and our gradient descent implementation in the first two benchmark domains, where the space is well-conditioned. On the other hand, in the LSI domain CMA-MEGA without the Adam implementation performed poorly. We conjecture that this is caused by the conditioning of the mapping from the latent space of the StyleGAN to the CLIP score.

Fig. 2 shows the QD-score values for increasing number of evaluations for each of the tested algorithms, with 95% confidence intervals. The figure also presents heatmaps of the CMA-MEGA (Adam) and the generated archive of the strongest QD competitor for each of the three domains. We provide generated archives of all algorithms in the supplemental material.

We visualize the top performing solutions in the LSI domain by uniformly sampling solutions from the archive of CMA-MEGA (Adam) and showing the generated faces in Fig. 3. We observe that as we move from the top right to the bottom left, the features matching the captions "a man with blue eyes" and "a person with red hair" become more prevalent. We note that these solutions were generated from a single run of CMA-MEGA (Adam) for 10,000 iterations.

Overall, these results show that using the gradient information in quality diversity optimization results in significant benefits in search efficiency, but adapting the gradient coefficients with CMA-ES is critical in achieving top performance.

7 Related Work

Quality Diversity. The precursor to QD algorithms [46] originated with diversity-driven algorithms as a branch of evolutionary computation. Novelty search [36], which maintains an archive of diverse solutions, ensures diversity though a provided metric function and was the first diversity-driven algorithm. Later, objectives were introduced as a quality metric resulting in the first QD algorithms: Novelty Search with Local Competition (NSLC) [37] and MAP-Elites [13, 40]. Since their inception, many works have improved the archives [17, 57, 53], the variation operators [56, 18, 11, 43], and the selection mechanisms [12, 52] of both NSLC and MAP-Elites. While the original QD algorithms were based on genetic algorithms, algorithms based on other derivative-free approaches such as evolution strategies [18, 10, 43, 11] and Bayesian optimization [34] have recently emerged.

Being stochastic derivative-free optimizers [8], QD algorithms are frequently applied to reinforcement learning (RL) problems [44, 3, 14] as derivative information must be estimated in RL. Naturally, approaches combining QD and RL have started to emerge [42, 9]. Unlike DQD, these approaches *estimate* the gradient of the reward function in action space and backpropagate this gradient through a neural network. Our proposed DQD problem differs by leveraging provided gradients for both the objective and measure functions.

Several works have proposed model-based QD algorithms. For example, the DDE-Elites algorithm [21] dynamically trains a variational auto-encoder (VAE) on the QD archive, then leverages the latent space of this VAE as a learned parameter space to optimize. The PoMS algorithm [48] builds on DDE-Elites by introducing a variation operator which samples based on the Jacobian of the learned VAE. These works differ by dynamically constructing a learned representation of the search space instead of leveraging the objective and measure gradients directly.

Latent Space Exploration. Several works have proposed a variety of methods for directly exploring the latent space of generative models. Methods on GANs include interpolation [55], gradient descent [4], importance sampling [59], and latent space walks [30]. Derivative-free optimization methods mostly consist of latent variable evolution (LVE) [5, 22], the method of optimizing latent space with an evolutionary algorithm. LVE was later applied to generating Mario levels [58] with targeted gameplay characteristics. Later work [19] proposed latent space illumination (LSI), the problem of exploring the latent space of a generative model with a QD algorithm. The method has only been applied to procedurally generating video game levels [19, 54, 51] and generating MNIST digits [60]. Follow-up work explored LSI on VAEs [50]. Our work improves LSI on domains where gradient information on the objective and measures is available with respect to model output.

8 Limitations and Future Work

Quality diversity (QD) is a rapidly emerging field [8] with applications including procedural content generation [25], damage recovery in robotics [13, 40], efficient aerodynamic shape design [20], and scenario generation in human-robot interaction [16]. We have introduced differentiable quality diversity (DQD), a special case of QD, where measure and objective functions are differentiable, and showed how gradient arborescence results in significant improvements in search efficiency.

As both MEGA variants are only first order differentiable optimizers, we expect them to have difficulty on highly ill-conditioned optimization problems. CMA-ES, as an approximate second order optimizer, retains a full-rank covariance matrix that approximates curvature information and is known to outperform quasi-Newton methods on highly ill-conditioned problems [23]. CMA-ME likely inherits these properties by leveraging the CMA-ES adaptation mechanisms and we expect it to have an advantage on ill-conditioned objective and measure functions.

While we found CMA-MEGA to be fairly robust to hyperparameter changes in the first two benchmark domains (linear projection, arm repertoire), small changes of the hyperparameters in the LSI domain led CMA-MEGA, as well as all the QD baselines, to stray too far from the mean of the latent space, which resulted in many artifacts and unrealistic images. One way to address this limitation is to constrain the search region to a hypersphere of radius $\sqrt{d}$, where d is the dimensionality of the latent space, as done in previous work [39].

While CLIP achieves state-of-the-art performance in classifying images based on visual concepts, the model does not measure abstract concepts. Ideally, we would like to specify "age" as a measure function and obtain quantitative estimates of age given an image of a person. We believe that the proposed work on the LSI domain will encourage future research on this topic, which we would in turn be able to integrate with DQD implementations to generate diverse, high quality content.

Many problems, currently modelled as optimization problems, may be fruitfully redefined as QD problems, including the training of deep neural networks. Our belief stems from recent works [49, 38], which reformulated deep learning as a multi-objective optimization problem. However, QD algorithms struggle with high-variance stochastic objectives and measures [32, 15], which naturally conflicts with minibatch training in stochastic gradient descent [6]. These challenges need to be addressed before DQD training of deep neural networks becomes tractable.

References

[1] Generating images from prompts using clip and stylegan. https://towardsdatascience.com/generating-images-from-prompts-using-clip-and-stylegan-1f9ed495ddda. MIT License.

[2] Youhei Akimoto, Yuichi Nagata, Isao Ono, and Shigenobu Kobayashi. Bidirectional relation between cma evolution strategies and natural evolution strategies. In *International Conference on Parallel Problem Solving from Nature*, pages 154–163. Springer, 2010.

[3] Kai Arulkumaran, Antoine Cully, and Julian Togelius. Alphastar: An evolutionary computation perspective. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 314–315, 2019.

[4] Piotr Bojanowski, Armand Joulin, David Lopez-Paz, and Arthur Szlam. Optimizing the latent space of generative networks. *arXiv preprint arXiv:1707.05776*, 2017.

[5] Philip Bontrager, Aditi Roy, Julian Togelius, Nasir Memon, and Arun Ross. Deepmasterprints: Generating masterprints for dictionary attacks via latent variable evolution. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–9. IEEE, 2018.

[6] Léon Bottou. Stochastic gradient descent tricks. In *Neural networks: Tricks of the trade*, pages 421–436. Springer, 2012.

[7] Joy Buolamwini and Timnit Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*, pages 77–91. PMLR, 2018.

[8] Konstantinos Chatzilygeroudis, Antoine Cully, Vassilis Vassiliades, and Jean-Baptiste Mouret. Quality-diversity optimization: a novel branch of stochastic optimization. *arXiv preprint arXiv:2012.04322*, 2020.

[9] Geoffrey Cideron, Thomas Pierrot, Nicolas Perrin, Karim Beguir, and Olivier Sigaud. Qd-rl: Efficient mixing of quality and diversity in reinforcement learning. *arXiv preprint arXiv:2006.08505*, 2020.

[10] Cédric Colas, Vashisht Madhavan, Joost Huizinga, and Jeff Clune. Scaling map-elites to deep neuroevolution. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, pages 67–75, 2020.

[11] Edoardo Conti, Vashisht Madhavan, Felipe Petroski Such, Joel Lehman, Kenneth O Stanley, and Jeff Clune. Improving exploration in evolution strategies for deep reinforcement learning via a population of novelty-seeking agents. *arXiv preprint arXiv:1712.06560*, 2017.

[12] Antoine Cully and Yiannis Demiris. Quality and diversity optimization: A unifying modular framework. *IEEE Transactions on Evolutionary Computation*, 22(2):245–259, 2017.

[13] Antoine Cully, Jeff Clune, Danesh Tarapore, and Jean-Baptiste Mouret. Robots that can adapt like animals. *Nature*, 521(7553):503, 2015.

[14] Adrien Ecoffet, Joost Huizinga, Joel Lehman, Kenneth O Stanley, and Jeff Clune. First return, then explore. *Nature*, 590(7847):580–586, 2021.

[15] Manon Flageat and Antoine Cully. Fast and stable map-elites in noisy domains using deep grids. In *Artificial Life Conference Proceedings*, pages 273–282. MIT Press, 2020.

[16] Matthew Fontaine and Stefanos Nikolaidis. A quality diversity approach to automatically generating human-robot interaction scenarios in shared autonomy. *Robotics: Science and Systems*, 2021. to appear.

[17] Matthew C Fontaine, Scott Lee, Lisa B Soros, Fernando de Mesentier Silva, Julian Togelius, and Amy K Hoover. Mapping hearthstone deck spaces through map-elites with sliding boundaries. In *Proceedings of The Genetic and Evolutionary Computation Conference*, pages 161–169, 2019.

[18] Matthew C Fontaine, Julian Togelius, Stefanos Nikolaidis, and Amy K Hoover. Covariance matrix adaptation for the rapid illumination of behavior space. In *Proceedings of the 2020 genetic and evolutionary computation conference*, pages 94–102, 2020.

[19] Matthew C Fontaine, Ruilin Liu, Julian Togelius, Amy K Hoover, and Stefanos Nikolaidis. Illuminating mario scenes in the latent space of a generative adversarial network. *In Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

[20] Adam Gaier, Alexander Asteroth, and Jean-Baptiste Mouret. Data-efficient design exploration through surrogate-assisted illumination. *Evolutionary computation*, 26(3):381–410, 2018.

[21] Adam Gaier, Alexander Asteroth, and Jean-Baptiste Mouret. Automating representation discovery with map-elites. *arXiv preprint arXiv:2003.04389*, 2020.

[22] Federico A Galatolo, Mario GCA Cimino, and Gigliola Vaglini. Generating images from caption and vice versa via clip-guided generative latent space search. *arXiv preprint arXiv:2102.01645*, 2021.

[23] Tobias Glasmachers and Oswin Krause. The hessian estimation evolution strategy. In *International Conference on Parallel Problem Solving from Nature*, pages 597–609. Springer, 2020.

[24] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2672–2680, 2014.

[25] Daniele Gravina, Ahmed Khalifa, Antonios Liapis, Julian Togelius, and Georgios N Yannakakis. Procedural content generation through quality diversity. In *2019 IEEE Conference on Games (CoG)*, pages 1–8. IEEE, 2019.

[26] David Güera and Edward J Delp. Deepfake video detection using recurrent neural networks. In *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 1–6. IEEE, 2018.

[27] N. Hansen, A. Auger, O. Mersmann, T. Tusar, and D. Brockhoff. Coco: A platform for comparing continuous optimizers in a black-box setting. 2016.

[28] Nikolaus Hansen. The cma evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772*, 2016.

[29] Nikolaus Hansen, Anne Auger, Raymond Ros, Steffen Finck, and Petr Pošík. Comparing results of 31 algorithms from the black-box optimization benchmarking bbob-2009. pages 1689–1696, 07 2010. doi: 10.1145/1830761.1830790.

[30] Ali Jahanian, Lucy Chai, and Phillip Isola. On the" steerability" of generative adversarial networks. *arXiv preprint arXiv:1907.07171*, 2019.

[31] Norman L Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous univariate distributions, volume 2*, volume 289. John wiley & sons, 1995.

[32] Niels Justesen, Sebastian Risi, and Jean-Baptiste Mouret. Map-elites for noisy domains by adaptive sampling. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 121–122, 2019.

[33] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4401–4410, 2019.

[34] Paul Kent and Juergen Branke. Bop-elites, a bayesian optimisation algorithm for quality-diversity search. *arXiv preprint arXiv:2005.04320*, 2020.

[35] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[36] Joel Lehman and Kenneth O Stanley. Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary computation*, 19(2):189–223, 2011.

[37] Joel Lehman and Kenneth O Stanley. Evolving a diversity of virtual creatures through novelty search and local competition. In *Proceedings of the 13th annual conference on Genetic and evolutionary computation*, pages 211–218, 2011.

[38] Suyun Liu and Luis Nunes Vicente. The stochastic multi-gradient algorithm for multi-objective optimization and its application to supervised machine learning. *Annals of Operations Research*, pages 1–30, 2021.

[39] Sachit Menon, Alexandru Damian, Shijia Hu, Nikhil Ravi, and Cynthia Rudin. Pulse: Self-supervised photo upsampling via latent space exploration of generative models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2437–2445, 2020.

[40] Jean-Baptiste Mouret and Jeff Clune. Illuminating search spaces by mapping elites. *arXiv preprint arXiv:1504.04909*, 2015.

[41] Richard M Murray, Zexiang Li, and S Shankar Sastry. *A mathematical introduction to robotic manipulation*. CRC press, 2017.

[42] Olle Nilsson and Antoine Cully. Policy gradient assisted map-elites. 2021.

[43] Jørgen Nordmoen, Eivind Samuelsen, Kai Olav Ellefsen, and Kyrre Glette. Dynamic mutation in map-elites for robotic repertoire generation. In *Artificial Life Conference Proceedings*, pages 598–605. MIT Press, 2018.

[44] Jack Parker-Holder, Aldo Pacchiano, Krzysztof Choromanski, and Stephen Roberts. Effective diversity in population-based reinforcement learning. *arXiv preprint arXiv:2002.00632*, 2020.

[45] Justin K Pugh, Lisa B Soros, Paul A Szerlip, and Kenneth O Stanley. Confronting the challenge of quality diversity. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pages 967–974, 2015.

[46] Justin K Pugh, Lisa B Soros, and Kenneth O Stanley. Quality diversity: A new frontier for evolutionary computation. *Frontiers in Robotics and AI*, 3:40, 2016.

[47] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. *arXiv preprint arXiv:2103.00020*, 2021.

[48] Nemanja Rakicevic, Antoine Cully, and Petar Kormushev. Policy manifold search: Exploring the manifold hypothesis for diversity-based neuroevolution. *arXiv preprint arXiv:2104.13424*, 2021.

[49] Michael Ruchte and Josif Grabocka. Efficient multi-objective optimization for deep learning. *arXiv preprint arXiv:2103.13392*, 2021.

[50] Anurag Sarkar and Seth Cooper. Generating and blending game levels via quality-diversity in the latent space of a variational autoencoder. *arXiv preprint arXiv:2102.12463*, 2021.

[51] Jacob Schrum, Vanessa Volz, and Sebastian Risi. Cppn2gan: Combining compositional pattern producing networks and gans for large-scale pattern generation. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, pages 139–147, 2020.

[52] Konstantinos Sfikas, Antonios Liapis, and Georgios N Yannakakis. Monte carlo elites: Quality-diversity selection as a multi-armed bandit problem. *arXiv preprint arXiv:2104.08781*, 2021.

[53] Davy Smith, Laurissa Tokarchuk, and Geraint Wiggins. Rapid phenotypic landscape exploration through hierarchical spatial partitioning. In *International conference on parallel problem solving from nature*, pages 911–920. Springer, 2016.

[54] Kirby Steckel and Jacob Schrum. Illuminating the space of beatable lode runner levels produced by various generative adversarial networks. *arXiv preprint arXiv:2101.07868*, 2021.

[55] Paul Upchurch, Jacob Gardner, Geoff Pleiss, Robert Pless, Noah Snavely, Kavita Bala, and Kilian Weinberger. Deep feature interpolation for image content changes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7064–7073, 2017.

[56] Vassilis Vassiliades and Jean-Baptiste Mouret. Discovering the elite hypervolume by leveraging interspecies correlation. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 149–156, 2018.

[57] Vassilis Vassiliades, Konstantinos Chatzilygeroudis, and Jean-Baptiste Mouret. Using centroidal voronoi tessellations to scale up the multi-dimensional archive of phenotypic elites algorithm. *arXiv preprint arXiv:1610.05729*, 2016.

[58] Vanessa Volz, Jacob Schrum, Jialin Liu, Simon M Lucas, Adam Smith, and Sebastian Risi. Evolving mario levels in the latent space of a deep convolutional generative adversarial network. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 221–228, 2018.

[59] Tom White. Sampling generative networks. *arXiv preprint arXiv:1609.04468*, 2016.

[60] Yulun Zhang, Bryon Tjanaka, Matthew C. Fontaine, and Stefanos Nikolaidis. Illuminating the latent space of an mnist gan. *pyribs.org*, 2021. URL `https://docs.pyribs.org/en/stable/tutorials/lsi_mnist.html`. MIT License.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] (see section 8)
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] We include a discussion in the supplemental material.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] Source code with instructions to reproduce the experiments is included in the supplemental material.

(b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] Details about hyperparameters are included in the supplemental material

(c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] We ran statistical tests to test significance and we also show 95% confidence intervals in Fig. 2. We include additional results with error bars in the supplemental material.

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] We included this information in the supplemental material.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

(a) If your work uses existing assets, did you cite the creators? [Yes] We cite the creators of the pyribs library for the implementation of the QD algorithms, the creator of the StyleGAN+CLIP source code that we used in the latent space illumination domain and the relevant papers for StyleGAN and CLIP models.

(b) Did you mention the license of the assets? [Yes] We mention the license (MIT License) in the related references.

(c) Did you include any new assets either in the supplemental material or as a URL? [Yes] We include additional results in the supplemental material.

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]