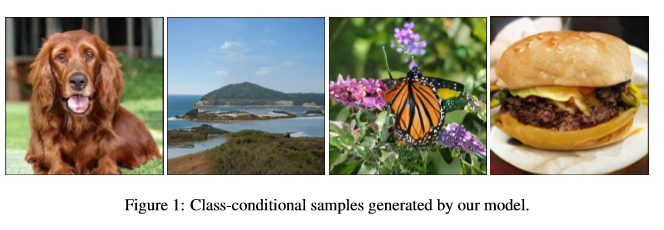
* 직교 정규화 (orthogonal regularization ) :
* Regularization: 직관적으로 생각하면, 어떤 문제를 해결하는 분류기를 찾고 싶을 때, 해당 분류기가 들어있는 통이 너무 커서 적절한 답을 찾기 힘들 때, 분류기들을 일단 채에 한번 걸른 후에 걸러진 것들 중에 답을 찾는 것과 비슷하다.
* 절단 트릭 (truncation trick)
* 인셉션 점수(IS) 와 프레쳇 인셉션 거리(FID): 두 점수 모두 GAN의 성능을 평가하기 위한 지표. IS는 높을수록 좋고, FID는 낮을수록 좋다.

<개요>

* 최근 생성 이미지 모델링의 발전에도 불구하고 ImageNet과 같은 복잡한 데이터 **세트에서 고해상도의 다양한 샘플을 성공적으로 생성하는 것은 어려운 목표**입니다. 이를 위해 우리는 아직 시도 된 최대 규모로 Generative Adversarial Networks를 교육하고 그러한 규모에 특정한 불안정성을 연구합니다. 우리는 생성기에 **직교 정규화**를 적용하면 간단한 "**절단 트릭**"을 이용할 수 있으며, 생성기 입력의 분산을 줄임으로써 **샘플 충실도**와 **다양성** 간의 절충을 **제어** 할 수 있습니다. 우리의 수정은 최첨단 조건부 이미지 합성의 새로운 상태를 설정하는 모델로 이어집니다. 128 × 128 해상도로 ImageNet을 교육 할 때, 당사 모델 (BigGAN)은 166.5의 초기 점수 (IS)와 7.4의 Fr'echet 초기 거리 (FID)를 달성하여 52.52의 이전 최고 IIS와 18.65의 FID를 향상시킵니다.



고해상도의 이미지를 생성하는 GAN에서 높은 점수를 얻게끔 시도하는 여러가지 방법.

1. 스케일링을 이용한 점수올리기 - 2-4 배 많은 매개 변수와 8 배 배치 크기로 모델을 훈련
2. 샘플의 다양성과 충실도 사이의 trade-off 를 제어하는 샘플링 기술인 truncation trick을 이용함
3. 대규모 GAN에 불안정성을 발견하고 그 불안정성을 새로운 기술과 기존 기술의 조합이 이러한 불안정성을 줄일 수 있지만 완전한 훈련 안정성은 극적인 성능 대 비용으로 만 달성 할 수 있음을 보여줌.

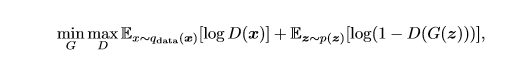
128 × 128 해상도로 ImageNet을 훈련 시킬 때, 당사의 모델 (BigGAN)은 IS를 (52.52->166.5)로 FID를 (18.65->7.4)로 향상시킵니다.

256x256에서는 IS를 232.5 로 FID를 8.1 로 향상

512x512에서는 IS를 241.5 로 FID를 11.5 로 향상

우리는 더 큰 모델의 데이터 세트 (JFT-300M)를 훈련시키고 우리의 디자인이 ImageNet에서 잘 전송되는지 입증합니다.

<기본지식 GAN>



여기서 의 z는 분포 p(z)에서 도출 된 잠재 변수

G와 D는 일반적으로 컨볼루션 신경망(CNN)이다. 보조 안정화 기술이 없으면이 교육 절차가 매우 취약하여 미세 조정 된 하이퍼 파라미터와 아키텍처 선택이 전혀 필요하지 않다.

G에서 스펙트럼 정규화 (Spectral Normalization)를 채택하면 안정성이 향상되어 반복 당 D 단계가 줄어든다는 사실을 발견

ProGan 은 일련의 증가하는 해상도에서 단일 모델을 훈련시켜 단일 클래스 환경에서 고해상도 GAN을 훈련

Miyato & Koyama (2018)에서 D는 피처와 생성 된 샘플을 구별하기위한 추가 증거로 피처와 학습 된 클래스 임베딩 간의 코사인 유사성을 사용하여 조건이 학습 된 클래스 프로토 타입과 일치하는 샘플 생성을 효과적으로 장려

Inception Score (IS, Salimans et al. (2016)) 및 Fr´echet Inception Distance (FID, Heusel et al. (2017))는 주목할만한 결함에도 불구하고 인기를 얻음

* 이 논문에서 그것들을 대략적인 샘플 품질 측정 수단으로 사용하고 이전 작업과 비교할 수 있다.

3 SCALING UP GANS

이 섹션에서는 GAN 교육을 확장하는 방법을 살펴보고 더 큰 모델과 큰 배치의 성능 이점을 얻습니다. SA-GAN 아키텍쳐를 사용 힌지 손실을 사용. 클래스 조건부 BatchNorm을 사용하여 클래스 정보를 제공하고 projection(투사)을 사용하여 D에 클래스 정보를 제공합니다. 최적화 설정은 수익률을 낮추고 G 단계마다 두개의 D 단계를 수행하는 수정을 사용하였다. 평가를 위해 Karras et al.에 따라 G의 가중치의 이동 평균을 사용합니다. 우리는 Orthogonal Initialization (Saxe et al., 2014)을 사용하는 반면, 이전 연구에서는 N (0,0.02I) (Radford et al., 2016) 또는 Xavier 초기화 (Glorot & Bengio, 2010)를 사용했습니다. 각 모델은 Google TPUv3 Pod (Google, 2018)의 128 ~ 512 코어에 대해 교육을 받았으며 일반적으로 기기별로가 아니라 모든 기기에서 G의 BatchNorm 통계를 계산합니다. 512x512 모델에서도 필요없는 점진적 성장 (Karras et al., 2018)을 찾습니다.

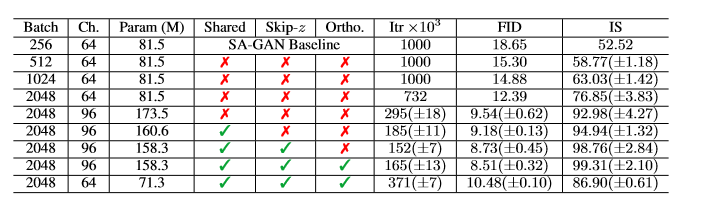


표 1.

Batch크기 / 총 매개변수의 수 / 레이어의 단위 수를 나타내는 채널 승수/ 공유 임베딩을 사용하는지/ 잠재된 레이어에서 여러 레이어로의 스킵연결을 사용하는지/ 직교 정규화를 사용하는지/ 106회 반복에 안정적인지 또는 반복에서 축소되는지

1~4행 이외의 결과는 8개의 임의 초기화에 의해서 계산된다.

우리는 기본 모델의 배치 크기를 늘리는 것으로 시작하여 즉시 큰 이점을 찾습니다. 표 1의 1-4 행은 배치 크기를 8 배로 늘리는 것만으로 최신 IS를 46 % 향상시킵니다. 이는 각 배치가 더 많은 모드를 포함하여 두 네트워크 모두에 더 나은 기울기(gradients)를 제공 한 결과라고 추측합니다. 이 스케일링의 주목할 만한 부작용 중 하나는 우리 모델이 더 나은 성능 침해에 도달하지만 완벽하게 훈련 붕괴 될 수 있다는 것입니다. 이 실험에서는 붕괴 직전에 저장된 체크 포인트 점수를 보고합니다. 그런 다음 각 레이어의 너비 (채널 수)를 50 % 늘려 두 모델의 매개 변수 수를 거의 두 배로 늘립니다. 이로 인해 IS의 21 %가 추가로 개선되어 데이터 세트의 복잡성에 비해 모델의 용량이 증가했습니다. 초기에 개선되지 않은 두 배의 개선은 다른 잔차 블록 구조를 사용하는 BigGAN 딥 모델에 적용되었습니다.

G의 조건부 BatchNorm 레이어에 사용되는( 클래스 포함) c에는 많은 가중치가 포함되어 있습니다. 각 임베딩마다 별도의 레이어를 사용하는 대신 각 레이어의 이익과 편향에 선형으로 투영되는 공유 임베딩을 사용하도록 선택합니다.

이는 계산 및 메모리 비용을 줄이고 훈련 속도 (주어진 성능에 도달하는 데 필요한 반복 횟수)를 37 % 향상시킵니다. 다음으로 노이즈 벡터 z에서 초기 레이어가 아닌 G의 여러 레이어에 직접 건너 뛰기 연결 (skip-z)을 추가합니다. 이 디자인의 직관은 G가 잠재 공간을 사용하여 다른 해상도와 계층 계층에서 기능에 직접 영향을 줄 수 있도록 하는 것입니다.

InBigGAN에서, 이것은 분해능(resolution, 서로 떨어져 있는 두 물체를 서로 구별할 수 있는 능력 ) 당 z를 하나의 청크로 분할하고 BatchNorm 이득 및 바이어스로 투영되는 조건부 벡터 c에 각 청크를 연결함으로써 완료됩니다. BigGAN-deep에서는 더 간단한 디자인을 사용하여 전체 z를 조건부 벡터로 묶어 덩어리로 나누지 않습니다. 이전 연구들 (Goodfellow etal., 2014; Denton et al., 2015)은이 개념의 변형을 고려했다; 우리의 구현은이 디자인의 작은 수정입니다. Skip-z는 약 4 %의 적당한 성능 향상을 제공하고 훈련 속도를 18 % 더 향상시킵니다.

TRADING OFF VARIETY AND FIDELITY WITH THE TRUNCATION TRICK

GAN은 임의의 이전 p (z)를 사용할 수 있지만 이전의 대부분의 작품은 z를 N (0, I) 또는 U [-1,1]에서 뽑아 내기로 선택했습니다. 우리는이 선택의 최적성에 의문을 가지고 부록 E의 대안을 탐색합니다.

**놀랍게도, 우리의 최상의 결과는 훈련에 사용 된 것과 다른 잠복 분포를 사용하여 샘플링 한 것입니다. z ∼ N (0, I)로 훈련 된 모델을 절단하고 잘린 법선에서 z를 샘플링하면 (범위를 벗어난 값이 해당 범위 내에 들어가도록 재 샘플링 됨) IS와 FID가 즉시 향상됩니다. 우리는 이것을 절단 트릭이라고 부릅니다.**

선택된 임계 값보다 큰 크기로 값을 resampling하여 z 벡터를 잘라 내면 전체 샘플 다양성의 감소 비용으로 개별 샘플 품질이 향상됩니다.

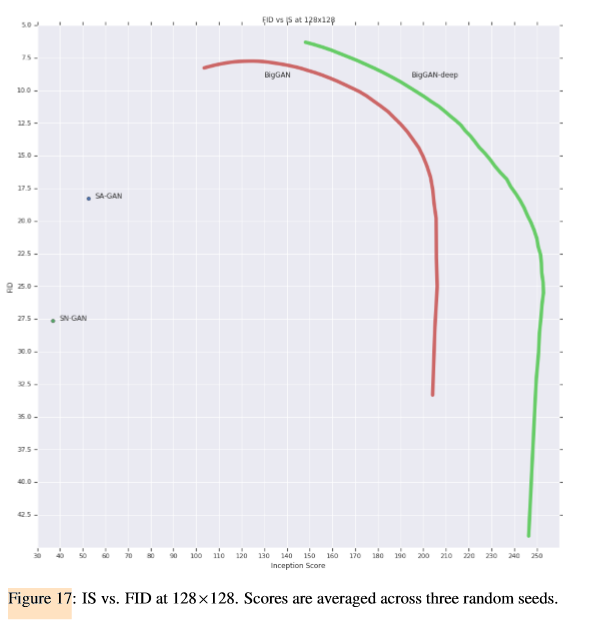


* 그림2. effects of increasing truncation 왼쪽에서 오른쪽으로 임계 값은 2, 1, 0.5, 0.04로 설정

임계 값이 감소하고 z의 요소가 0 (잠재 분포 모드)으로 잘 리면 개별 표본이 G의 출력 분포 모드에 접근합니다.

0에 가까울수록 그림이 같아진다. -> 정확도 올라감 -> 다양성 떨어짐 ㅠㅠ -> 이것이 trade-off

이 기법을 사용하면 주어진 G에 대해 샘플 품질과 다양성 사이의 균형 잡힌 사후 선택을 가능하게합니다. 특히, 임계 값 범위에 대해 FID와 IS를 계산하여 정밀도를 연상시키는 다양성 충실도 곡선을 얻을 수 있습니다



IS는 클래스 조건부 모델의 다양성 부족을 처벌하지 않기 때문에 잘림 임계 값을 줄이면 IS가 직접 증가합니다 (정밀도와 유사). FID는 다양성의 부족을 불이익을 주지만 (정확한 리콜) 정밀도를 보상하므로 처음에는 FID가 약간 개선되지만, 잘림이 0에 가까워지고 다양성이 줄어들면 FID가 급격히 떨어집니다. 훈련에서 볼 수있는 것과 다른 잠복으로 샘플링하여 발생하는 분포 변화는 많은 모델에서 문제가됩니다. 우리의 더 큰 모델 중 일부는 잘릴 수 없으며, 절단 된 잡음이 공급 될 때 포화 인공물 (그림 2 (b))을 생성합니다.

 그림 2(b)

이 문제를 해결하기 위해, z의 전체 공간이 양호한 출력 샘플에 매핑되도록 Gtobe를 부드럽게 조정하여 강제로 잘림을 강화합니다. 이를 위해 직교 정규화 (Brock et al., 2017)를 사용하여 직교 조건을 직접 시행합니다.



여기서 W는 가중치 행렬이고 β는 하이퍼 파라미터입니다. 이 정규화는 종종 너무 제한적인 것으로 알려져 있으므로 모델에 원하는 평활도를 부여하면서 제약 조건을 완화하도록 설계된 여러 변형을 탐색합니다. 우리가 가장 잘 작동하는 버전은 정규화에서 대각선 용어를 제거하고 필터 간의 쌍별 코사인 유사성을 최소화하지는 않지만 표준을 제한하지는 않습니다.



여기서 1은 모든 요소가 1로 설정된 행렬을 나타냅니다. 우리는 β 값을 스윕하고 10-4를 선택하여이 작은 추가 페널티를 찾아 모델이 잘릴 수있는 가능성을 향상시킵니다. 표 1에 걸쳐서, 직교 정규화가없는 경우, 모델의 16 %만이 절단이 가능하고, 직교 정규화로 훈련 된 경우 60 %에 해당하는 것으로 관찰되었다.

요약: 현재의 GAN 기술은 규모가 큰 모델과 분산 된 대규모 배치 교육을 수행하기에 충분

이러한 개선에도 불구하고 우리 모델은 훈련 붕괴를 겪으며 실제로 조기에 정지해야합니다. 다음 두 섹션에서 우리는 이전 작업에서 안정적인 설정이 대규모로 적용될 때 불안정 해지는 이유를 조사합니다.

여기까지는 나도 이해함.

근데 이 다음부터는 너무 어렵다.

4.1 CHARACTERIZING INSTABILITY: THE GENERATOR

이유를 생성기에서 찾는듯.

이전의 많은 연구는 다양한 분석 각도와 장난감 문제에서 GAN 안정성을 조사했지만, 불안정한 소규모 설정에서는 발생하는 불안정성이 발생하여 직접적인 분석이 필요합니다. 훈련 중 무게, 기울기 및 손실 통계를 모니터링합니다 (Odena et al., 2018)와 유사하게 훈련 붕괴가 시작될 수있는 지표를 검색합니다. 각 가중치 행렬의 상위 3 개의 특이 값 σ0, σ1, σ2가 가장 유익한 것으로 나타났습니다. Alrnoldi 반복 방법 (Golub & der Vorst, 2000)을 사용하여 효율적으로 계산할 수 있습니다.이 방법은 전력 반복 방법을 추가 특이 벡터 및 값의 추정으로 확장합니다.

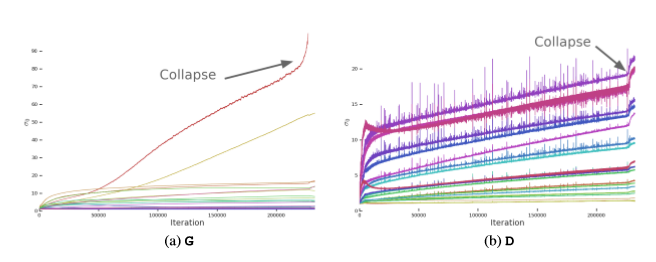
그림 3.스펙트럼 정규화 전 G (a)와 D (b)의 레이어에서 첫 번째 특이 값 σ0의 일반적인 도표. G의 대부분의 레이어는 잘 동작하는 스펙트럼을 갖지만 제약이 없으면 작은 서브 세트가 훈련 전체에서 커지고 붕괴시 폭발합니다. D의 스펙트럼은 더 시끄럽지 만 더 잘 작동합니다. 빨간색에서 보라색까지의 색상은 깊이가 증가 함을 나타냅니다.

그림 3 (a)와 부록 F에서 볼 수 있듯이 명확한 패턴이 나타납니다. 대부분의 G 층은 훈련 기간 동안 성장하고 붕괴시 폭발하는 스펙트럼 규범을 가진 잘 작동하는 스펙트럼 규범을 가지고 있습니다. 그러나 일부 레이어는 잘못 작동합니다. 일반적으로 G의 첫 번째 레이어는 과도하고 완전하지 않습니다

이 병리가 붕괴의 원인인지 또는 단순히 증상인지 확인하기 위해, 우리는 스펙트럼 폭발을 명시 적으로 방지하기 위해 G에 추가 컨디셔닝을 적용하는 효과를 연구합니다. 먼저, 각 가중치의 상위 특이 값 σ0을 고정 값 σreg 또는 두 번째 특이 값 r · sg (σ1)의 일부 비율 r로 직접 정규화합니다. σ1 증가). 대안 적으로, 우리는 대신에 σ0을 클램핑하기 위해 부분 특이 값 분해를 사용한다. 가중치 W, 첫 번째 특이 벡터 u0 및 v0, σ0이 클램핑 될 값을 클램핑하면 가중치가 다음과 같이됩니다.



Spectral Normalization의 유무에 관계없이 이러한 기술은 σ0 또는 σ0 σ1의 점진적인 증가 및 폭발을 방지하는 효과가 있지만, 경우에 따라 성능이 약간 향상되지만 조합은 훈련 붕괴를 방지하지 않습니다. 이 증거는 **투자 가능성을 조절하는 동안 안정성을 보장하는 데 충분하지 않다는 것을 암시합니다. 따라서 우리는 D로 관심을 돌립니다.**

**결론: 안정성 보장 안됨^^**

4.2 CHARACTERIZING INSTABILITY: THE DISCRIMINATOR

G와 마찬가지로 D의 가중치 스펙트럼을 분석하여 동작에 대한 통찰력을 얻은 다음 추가 제약 조건을 적용하여 훈련을 안정화시킵니다. 그림 3 (b)는 D에 대한 전형적인 σ0 플롯을 보여줍니다 (부록 F에 추가 플롯이 있음). G와는 달리, 스펙트럼은 노이즈가 있고 σ0/σ1은 잘 작동하며 특이 값은 훈련 기간 동안 커지지 만 폭발하지 않고 붕괴 될 때만 증가합니다.

D의 스펙트럼이 급등하면 주기적으로 매우 큰 기울기가 수신 될 수 있지만 Frobenius 규범이 부드럽다는 것을 알 수 있습니다 (부록 F).이 효과는 주로 몇 가지 단수 방향에 집중되어 있음을 나타냅니다. 우리는이 소음이 적대적 훈련 과정을 통한 최적화의 결과라고 결론 지었다. 여기서 G는 주기적으로 D를 강하게 방해하는 배치를 생성한다. 이 스펙트럼 노이즈가 불안정성과 인과 관계가있는 경우, 자연스러운 카운터는 그라디언트 페널티를 사용하여 D의 Jacobian의 변경을 명시 적으로 규칙 화하는 것입니다. 우리는 Mescheder 등의 R1 제로 중심 그라디언트 페널티를 탐구합니다. (2018) :



기본적으로 제안 된 γ 강도가 10이면 훈련이 안정되고 G와 D에서 스펙트럼의 부드러움과 경계가 향상되지만 성능이 크게 저하되어 IS가 45 % 감소합니다. 페널티 강도가 1 (급격한 붕괴가 발생하지 않는 가장 낮은 강도)로 감소하더라도 IS는 20 % 감소합니다.

Orthogonal Regularization, DropOut (Srivastava et al., 2014) 및 L2 (자세한 내용은 부록 I 참조)의 다양한 강점으로이 실험을 반복하면 이러한 정규화 전략에 대해 유사한 동작이 나타납니다. 하지만 상당한 성능 비용이 발생합니다.

D는 실습 과정에서 제로화가 사라지지만, 급격한 점핑 붕괴 (AppendixF)에 따른다는 점을 잘 알고있다. D의 암기 (Gulrajani et al. (2017)와 관련된)에 대한 간단한 테스트로서, 우리는 이미지 넷 트레이닝 및 검증 세트에 따라 미분류 판별자를 평가하고, 샘플의 비율이 실제로 생성되거나 생성 된 비율을 측정합니다. 훈련 정확도는 일관되게 98 % 이상이지만, 검증 정확도는 50-55 %의 범위에 속하며 (정규화 전략에 관계없이) 무작위 추측보다 나쁘지 않습니다. 이것은 D가 실제로 훈련 세트를 암기하고 있음을 확인합니다. 우리는 이것을 명시 적으로 일반화하는 것이 아니라 훈련 데이터를 증류하고 G에 유용한 학습 신호를 제공하는 D의 역할과 일치한다고 생각합니다. 추가 실험과 논의는 부록 G에 제공됩니다.

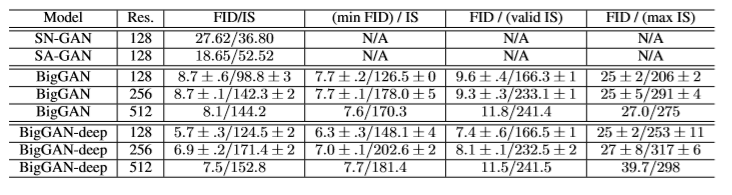
요약:

안정성은 G 또는 D에서만이 아니라 적대적 훈련 과정을 통한 상호 작용에서 비롯된 것입니다. 열악한 컨디셔닝의 증상은 불안정성을 추적하고 식별하는 데 사용될 수 있지만, 트레이닝에 필요한 합리적인 컨디셔닝을 보장하여 최종 트레이닝 붕괴를 방지하기에 충분합니다. D를 강력하게 제한하여 안정성을 강화할 수 있지만 그렇게하면 성능이 크게 저하됩니다. 현재의 기술을 사용하면이 컨디셔닝을 완화하고 훈련 후반 단계에서 붕괴가 발생하여 더 나은 최종 성능을 달성 할 수 있으며,이 때 모델은 충분한 결과를 얻도록 훈련됩니다.

5 EXPERIMENTS

5.1 EVALUATION ON IMAGENET

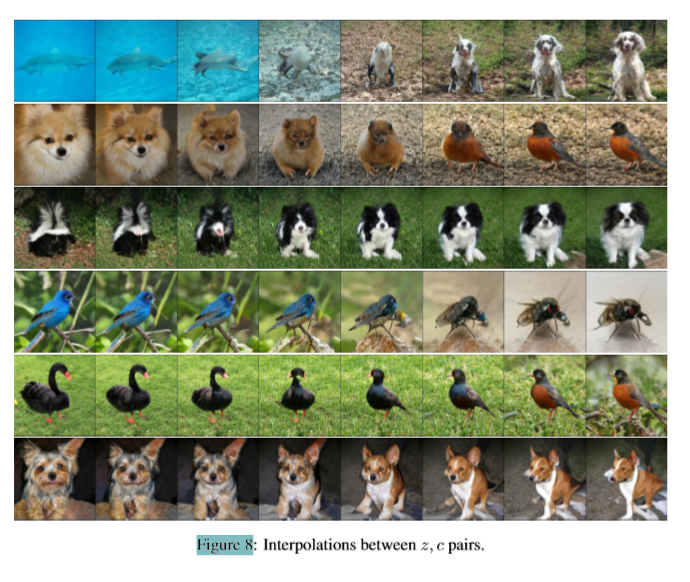
ImageNet ILSVRC 2012 (Russakovsky et al., 2015)에 대한 모델을 표 1, 행 8의 설정을 사용하여 128x128, 256x256 및 512x512 해상도로 평가합니다.

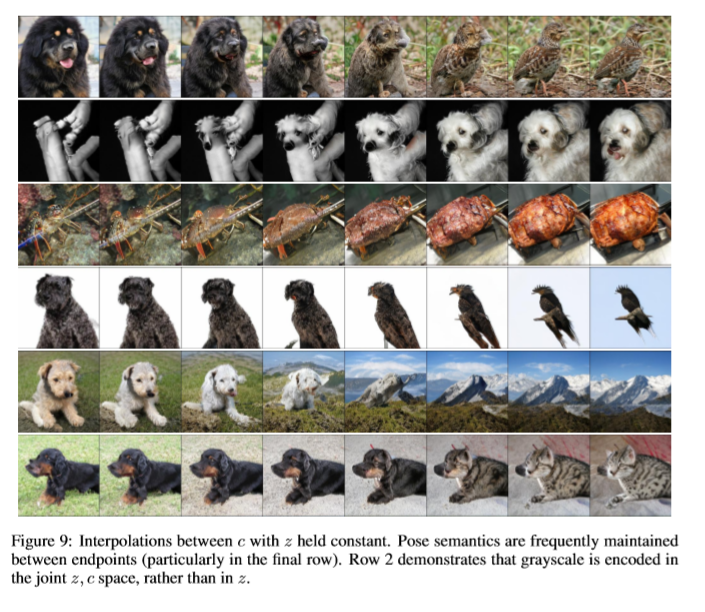
표 2 : 다른 해상도의 모델 평가. 우리는 잘리지 않은 점수 (3 열), 최고 FID (4 열), 검증 데이터의 IS (5 열) 및 maxIS (6 열)의 점수를 기록합니다. 표준 편차는 계산 된 세 가지 무작위 초기화.

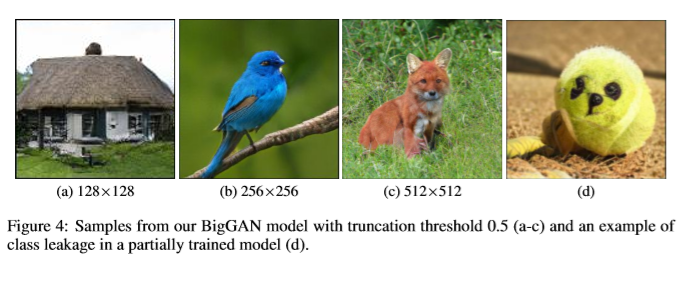
우리는 IS2와 FIDinTable2를보고합니다. 우리의 모델은 다양한 샘플의 품질을 평가할 수 있지만, 종래 기술과 비교할 수있는 최선의 방법은 불분명합니다. 그에 따라 부록 D에 완전한 곡선과 함께 세 가지 설정으로 값을보고합니다. 먼저, 최고의 FID를 달성하는 잘림 설정에서 FID / IS 값을보고합니다. 둘째, 우리는 모델의 IS가 실제 검증 데이터에 의해 달성되는 절단 설정에서 FID를보고합니다. 셋째, 우리는 품질을 극대화하기 위해 얼마나 많은 품종을 교환해야하는지 보여주기 위해 각 모델이 달성 한 최대 IS에서 FID를보고합니다. 세 가지 경우 모두, 당사의 모델은 Miyato et al.에 의해 달성 된 이전의 최신 IS 및 FID 점수를 능가합니다.

BigGAN 모델에 추가하여 종이의 첫 번째 버전을 소개하고 많은 실험을 사용합니다 (달리 명시되지 않은 경우). 잔차 블록의 다른 구성을 사용하는 4 배 더 깊은 모델 (BigGAN- 깊이)도 제시합니다. 표 2에서 볼 수 있듯이 BigGAN-deep은 모든 해상도 및 지표에서 BigGAN보다 성능이 뛰어납니다. 이것은 우리의 발견이 다른 아키텍처로 확장되고 샘플 품질의 심층적 인 향상을 증가 시킨다는 것을 확인합니다.

모델의 샘플 품질과 결합 된 D가 훈련 세트를 능가한다는 관찰은 단순한 기억력의 훈련 포인트가 아닌지에 대한 명백한 의문을 제기합니다. 이를 테스트하기 위해, 픽셀 공간과 사전 훈련 된 분류 네트워크 (Appendix A)의 특징 공간에서 클래스 단위로 가장 가까운 이웃 분석을 수행했습니다. 또한, 그림 8과 9에서 샘플과 클래스 간 보간 (z는 일정하게 유지됨) 사이의 보간을 나타냅니다. 우리의 모델이 단순히 훈련 데이터를 암기하는 것이 아니라고 제안하면서 샘플에 대한 가장 가까운 이웃은 시각적으로 명확하지 않습니다.

그림 8.

그림 9. z와 c 사이의 보간은 일정하게 유지됩니다. 포즈 의미론은 끝점 사이에서 (특히 최종 행) 자주 유지됩니다. 행 2는 그레이 스케일이 z가 아닌 조인트 z, c 공간에서 인코딩됨을 보여줍니다.

그림 4. 절단 임계 값 0.5 (a-c) 및 부분적으로 훈련 된 모델 (d)에서 클래스 누출의 예가있는 BigGAN 모델의 샘플.

부분적으로 훈련 된 모델의 일부 고장 모드는 이전에 관찰 된 것과 다릅니다. 대부분의 이전 실패에는 로컬 아티팩트 (Odena et al., 2016), 객체 대신 텍스처 Blob으로 구성된 이미지 (Salimans et al., 2016) 또는 표준 모드 축소가 포함됩니다. 그림 4 (d)에 의해 예시 된 바와 같이, 한 클래스의 이미지가 다른 클래스의 속성을 포함하는 클래스 누출을 관찰합니다. 또한 ImageNet의 많은 클래스가 다른 모델보다 더 어렵다는 것을 발견했습니다. 우리 모델은 군중 (데이터 집합의 작은 부분을 구성하고 더 큰 규모의 구조를 가짐)보다 개 (데이터 집합의 많은 부분을 구성하고 대부분 텍스처로 구별됨)를 생성하는 데 더 성공적입니다.

5.2 ADDITIONAL EVALUATION ON JFT-300M

우리의 설계 선택이 더 크고 더 복잡하고 다양한 데이터 세트에 효과적임을 확인하기 위해 JFT-300M의 하위 집합에 대한 시스템 결과도 제시합니다 (Sun et al., 2017). 전체 JFT-300M 데이터 세트에는 18K 범주로 레이블 된 300M 실제 이미지가 포함되어 있습니다. 카테고리 분포는 매우 길기 때문에 가장 일반적인 레이블이 8.5K 인 이미지 만 유지하기 위해 데이터 세트를 서브 샘플링합니다. 결과 데이터 집합에는 ImageNet보다 2 배 큰 292M 이미지가 포함됩니다. 라벨이 여러 개인 이미지의 경우 이미지가 샘플링 될 때마다 단일 라벨을 무작위로 독립적으로 샘플링합니다. 이 데이터 세트에 대해 훈련 된 GAN에 대한 IS 및 FID를 계산하기 위해이 데이터 세트에 대해 훈련 된 Inception v2 분류기 (Szegedy et al., 2016)를 사용합니다. 정량 결과는 표 3에 제시되어있다.

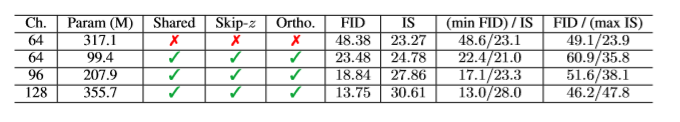


표 3. JFT-300M 검증 세트의 이미지는 IS는 50.88이고 FID는 1.94입니다.

모든 모델은 배치 크기 2048로 훈련됩니다. SA-GAN (Zhang et al., 2018)과 유사하지만 배치 크기가 더 큰 모델의 비교 버전을 ImageNet에서 최상의 결과를 얻기 위해 적용된 모든 기술을 사용하는 "전체"BigGAN 모델과 비교했습니다 ( 공유 임베딩, 스킵 -Z 및 직교 정규화). 우리의 결과는이 기술이 동일한 모델 용량 (64 개의 기본 채널)에서 더 큰 데이터 세트의 설정에서도 실질적으로 성능을 향상 시킨다는 것을 보여줍니다. 또한이 규모의 데이터 세트의 경우 128 개 기본 채널의 용량에서 4 개의 모델을 확장하여 상당한 개선을 볼 수 있으며 ImageNet GAN의 경우 추가 용량이 유리하지 않은 것으로 나타났습니다.

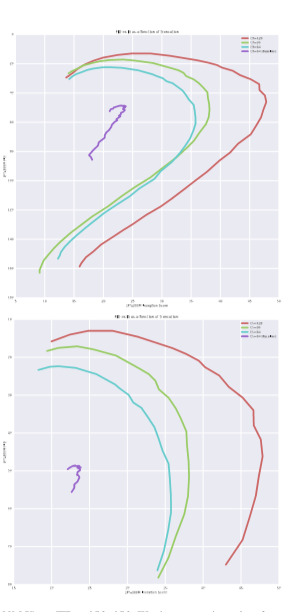
그림 19

그림 19 (부록 D)에는이 데이터 세트에 대해 훈련 된 모델에 대한 절단 도표가 나와 있습니다. 잘림 한계가 σ ≈ 0으로 가장 높은 충실도 점수를 생성하는 경향이있는 ImageNet과는 달리, 잘림 값 σ가 0.5에서 1 사이 일 때 IS는 일반적으로 JFT-300M 모델에 대해 최대화됩니다. 이는 JFT-300M 라벨의 클래스 내 가변성 및 다양한 크기의 여러 객체가 포함 된 이미지를 포함하는 이미지 분포의 열적 복잡성으로 인한 것입니다. 흥미롭게도 훈련이 많이 정규화되지 않고 붕괴되는 경향이있는 ImageNet에서 훈련 된 모델과는 달리 (섹션 4), JFT-300M에서 훈련 된 모델은 수십만 회 이상 안정적으로 유지됩니다. 이는 ImageNet을 넘어 더 큰 데이터 세트로 이동하면 GAN 안정성 문제를 부분적으로 완화 할 수 있음을 나타냅니다.

기본 모델을 변경하지 않고이 데이터 세트에서 달성 한 기본 GAN 모델에 대한 개선은 확장 된 용량을 넘어서서 표준화 된 기술을 훈련하지 않고 우리의 발견이 ImageNet에서 확장 성과 복잡성을 가진 데이터 세트로 확장되어 지금까지 생성되지 않은 이미지 모델에는 전례가없는 것을 보여줍니다.

결론:

우리는 Generative Adversarial Networks가 충실도 및 생성 된 샘플의 다양성 측면에서 스케일 업을 통해 여러 범주의 자연 이미지를 모델링하도록 훈련되었음을 입증했습니다. 그 결과, 당사의 모델은 ImageNet GAN 모델간에 새로운 수준의 성능을 설정하여 최신 기술을 크게 향상 시켰습니다. 우리는 또한 대규모 GAN의 훈련 행동에 대한 분석을 제시하고, 무게의 단일 값으로 안정성을 특성화하고 안정성과 성능 사이의 상호 작용에 대해 논의했습니다.