Title	Fake data - the future of advanced computer vision systems
Author(s)	Corcoran, Peter
Publication Date	2018-10-22
Publication Information	Corcoran, Peter. (2018). Fake data - the future of advanced computer vision systems. Plenary paper presented at the 2018 Duke Kunshan University International Innovation Symposium on Smart Camera Technologies, Duke Kunshan University Campus, Shanghai, China, 22-23 November, DOI: 10.13025/qbe4-qr97
Publisher	NUI Galway
Link to publisher's version	https://dx.doi.org/10.13025/qbe4-qr97
Item record	http://hdl.handle.net/10379/15508

Fake Data – the Future of Advanced Computer Vision Systems

Cognitive, Connected & Computational Imaging

Peter Corcoran

18th Nov 2018

ACKNOWLEDGEMENTS

- Portions of this research were funded under the SFI Strategic Partnership Program by Science Foundation Ireland (SFI) and FotoNation Ltd.
    - Project ID: 13/SPP/I2868; *Next Generation Imaging for Smartphone and Embedded Platforms*.

- This work was also supported by an Irish Research Council Employment Based PhD Award.
    - Project ID: EBPPG/2016/280.

- Funding and direction for much of this research was also provided by FotoNation Ltd. (Ireland) and its parent corporation Xperi Inc. (US).

OVERVIEW OF TODAY'S TALK

- **Topic #1 – The Connection between Deep Learning & Computer Vision**
  - **Some Background on how I fell into the Deep Learning Pool**
- Topic #2 – The Importance of Data for Deep Learning
  - What is Data Augmentation & why we need it
  - Smart-Augmentation – learning new ways to augment datasets
  - Boundary Equilibrius GAN (BEGAN) – a GAN to generate Face Images
  - Latent Spaces, learning Annotations & generating Custom Data
- Topic #3 – Virtual Reality as a source of Datasets
  - Real Data is difficult …
  - Fake data is better … and the tools are here today!



NEXT GENERATION IMAGING FOR SMARTPHONES

SFI Partnership Project (2015-2019)

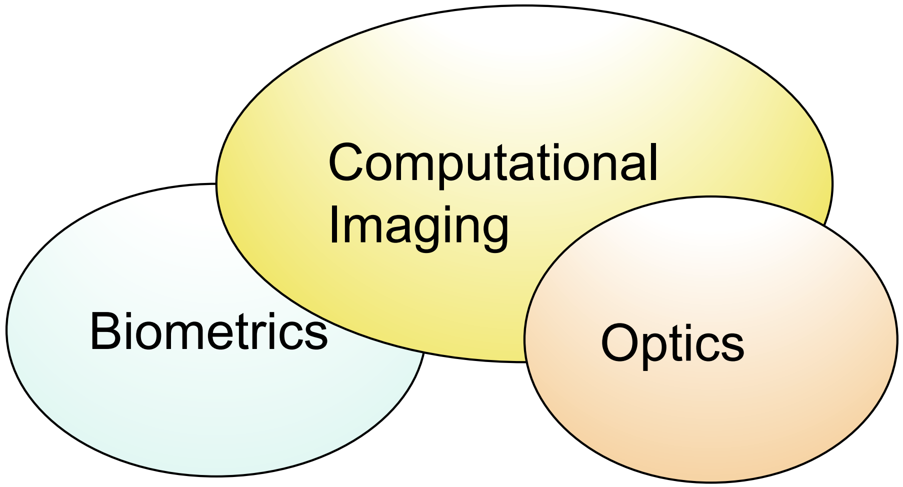
8 PhD Researchers & 2.5 Postdocs

NUI Galway
OÉ Gaillimh

C3Imaging Centre, College of Engineering & Informatics,
National University of Ireland Galway

SCOPE & OVERVIEW OF ORIGINAL PROPOSAL *(CONCEIVED IN 2014)*

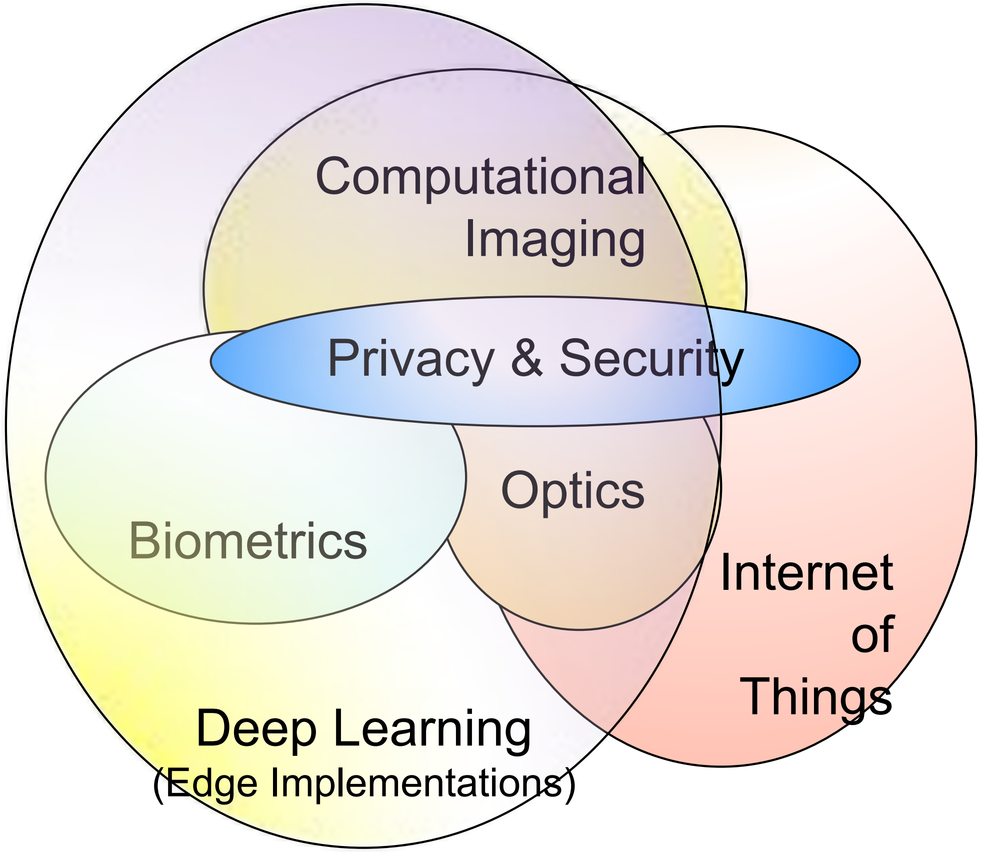
- Three Main Focus Areas for Research
  - Biometrics
  - Computational Imaging
  - Optics



- Additional Goals
  - Build a 'team' of ESR's covering these areas
  - Execute research with significant societal impact
  - Achieve a critical mass of expertise and industry support to establish a research centre in this field

TODAY …

- "DL at the Edge" overlaps all of these research areas;
- New 'Foundation Technologies' enhance data curation and improved AI Networks accuracy & performance beyond SoA;
- Parallel focus occuring in the Industry Partner, Xperi;



TOPIC #2 THE IMPORTANCE OF DATA

- Topic #1 – The Connection between Deep Learning & Computer Vision
  - Some Background on how I fell into the Deep Learning Pool
- **Topic #2 – The Importance of Data for Deep Learning**
  - What is Data Augmentation & why we need it
  - Smart-Augmentation – learning new ways to augment datasets
  - Boundary Equilibrius GAN (BEGAN) – a GAN to generate Face Images
  - Latent Spaces, learning Annotations & generating Custom Data
- Topic #3 – Virtual Reality as a source of Datasets
  - Real Data is difficult …
  - Fake data is better … and the tools are here today!

embedded VISION SUMMIT 2018

Getting More from Your Datasets: Data Augmentation, Annotation and Generative Techniques

Peter Corcoran, Joseph Lemley & Shabab Bazrafkan

18th May 2018

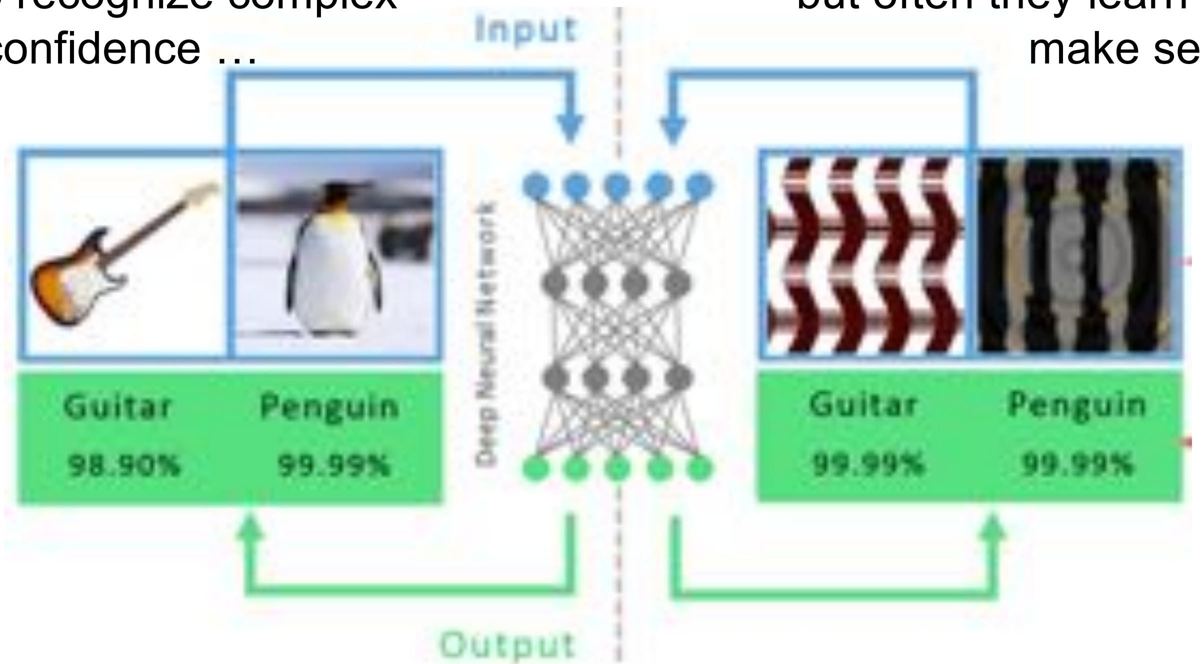
C3Imaging Centre, College of Engineering & Informatics, National University of Ireland Galway

Some Observations on Deep Neural Networks

DNNS ARE SMART, BUT NOT ALWAYS…

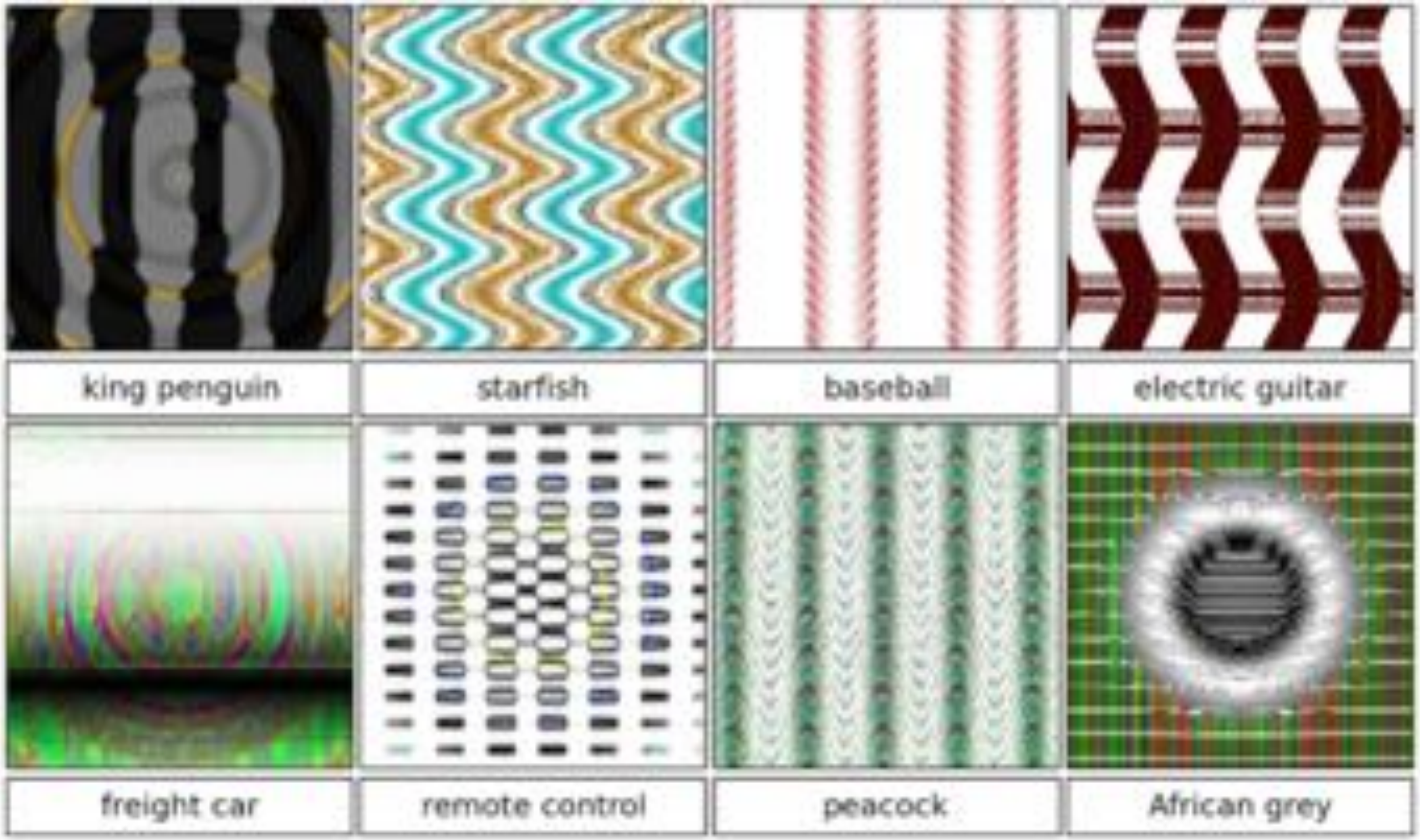
DNNs can learn to recognize complex objects with high confidence …

but often they learn features that don’t make sense to a human …

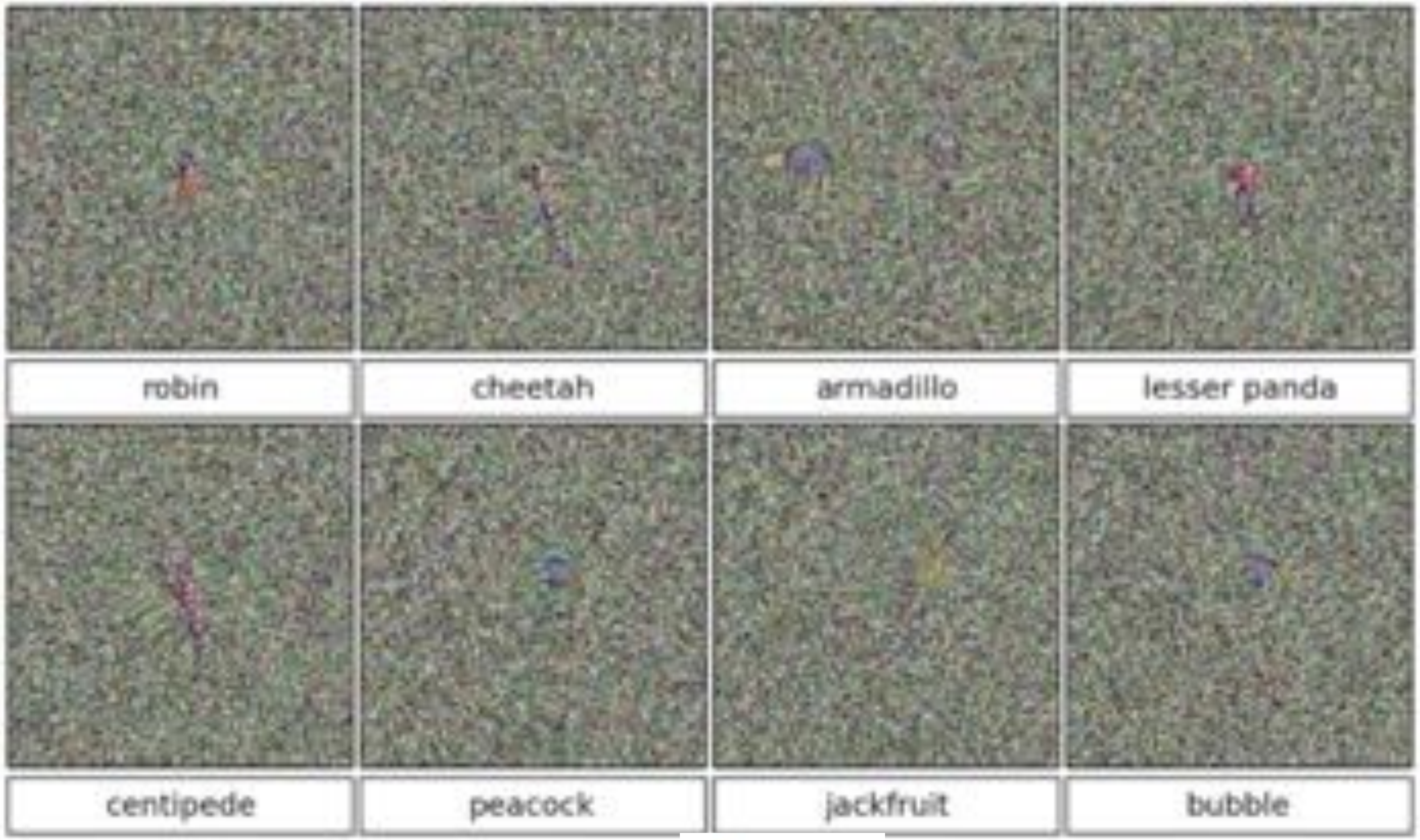


- Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images - ***Nguyen, et al - 2014***

DEEP NETWORKS CAN BE FOOLED BY CERTAIN 'LEARNED' (ADVERSARIAL) PATTERNS …



IN FACT "NOISE" WILL SOMETIMES WORK …



THE 1ST **MAIN POINT** FOR TODAY!

- Deep Neural Networks can appear very clever, but ….

- …. they are only as good as the Data that is used to Train them…

TOPIC #2.1 DATA AUGMENTATION

- Topic #1 – The Connection between Deep Learning & Computer Vision
  – Some Background on how I fell into the Deep Learning Pool
- **Topic #2 – The Importance of Data for Deep Learning**
  – **What is Data Augmentation & why we need it**
  – Smart-Augmentation – learning new ways to augment datasets
  – Boundary Equilibrius GAN (BEGAN) – a GAN to generate Face Images
  – Latent Spaces, learning Annotations & generating Custom Data
- Topic #3 – Virtual Reality as a source of Datasets
  – Real Data is difficult …
  – Fake data is better … and the tools are here today!

BASIC AUGMENTATION - ZOOMING

Augmenting small datasets is important and challenging. You are not adding much new information into the network, but by augmenting the data you are training the network not to overfit your dataset with regards to the type of augmentation.

Augmentation helps ensure that your Network learns ***semantically correct*** features!

BASIC AUGMENTATION - ROTATION

In an image classification task (e.g. dog/cat binary classification), if you rotate the image in various angles you are training the network to be invariant to rotation of the objects in the images.

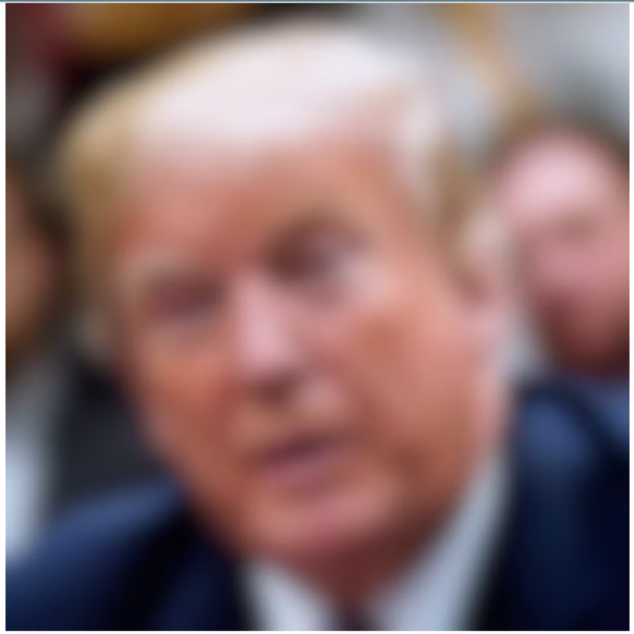
AUGMENTATION NEEDS TO BE APPLIED TO EACH DATA CLASS

BASIC AUGMENTATION – GAUSSIAN BLUR

R = 5

R = 15

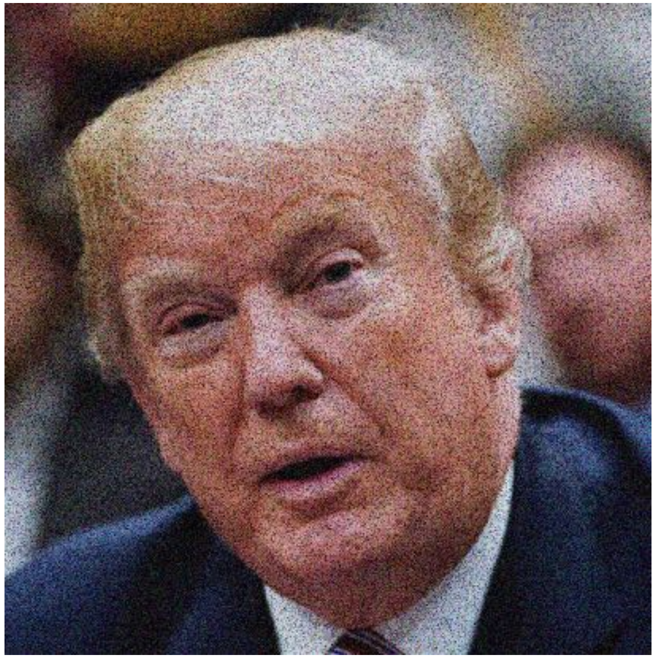
R = 30



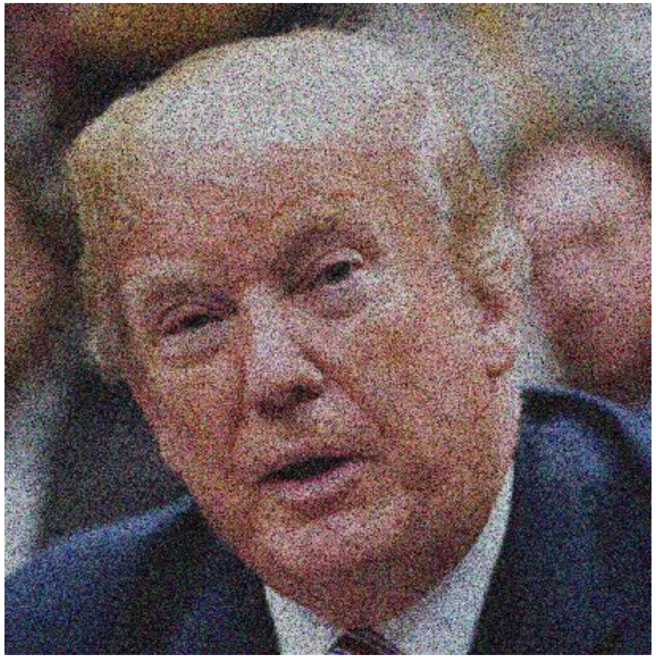
So although new “authentic” information isn't added into the network the “synthetic” data augmentation added into the network can both improve the results attained from the network and allow for training with less data.

BASIC AUGMENTATION – NOISE

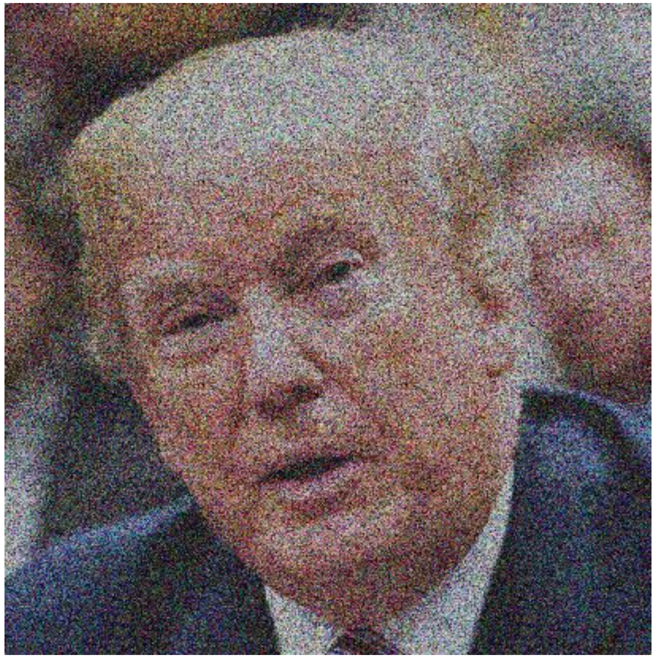
Var = 40



Var = 70



Var = 100



THE PURPOSE OF AUGMENTATION

- The Main Goal of Augmentation to prevent the deep network (DNN) from ‘learning’ the wrong features (*overfitting*)
    - Here we considered a simple classification problem but augmentation is a valuable tool in other DNN contexts …

THE 2ND **MAIN POINT** FOR TODAY!

- "Good Data" is needed to train a Deep Neural Network, but ….

- …. it needs to be 'Augmented' or the Network can learn some *wrong* characteristics of the data …

THE NATURE OF AUGMENTATION DEPENDS ON THE UNDERLYING PROBLEM

- Consider iris authentication on a smartphone
  - Iris resolution is relatively small (<100 pixels)
  - Smartphone is handheld (image blur)
  - Lighting conditions can vary significantly
- A key challenge is to accurately *segment* the iris regions
- High quality iris image datasets with many samples are available ….
  - …. but *poor quality images* needed for training …

HIGH QUALITY IRIS DATASETS USED TO DEVELOP TRAINING DATA & **GROUND TRUTH**

Iris diameter = 300+ pixels; **ground truth** *(determined from high quality commercial algorithms)* is at https://goo.gl/JVkSyG

Bath 800

CASIA 1000

IRIS FOR MOBILE

TYPICAL DATA IS LOW QUALITY FROM MOBILE HANDSETS

Iris diameter < 100 pixels; augmentation code available at:
https://github.com/C3Imaging/Deep-Learning-Techniques/blob/Iris_SegNet/DBaugmentation/DBaug.m

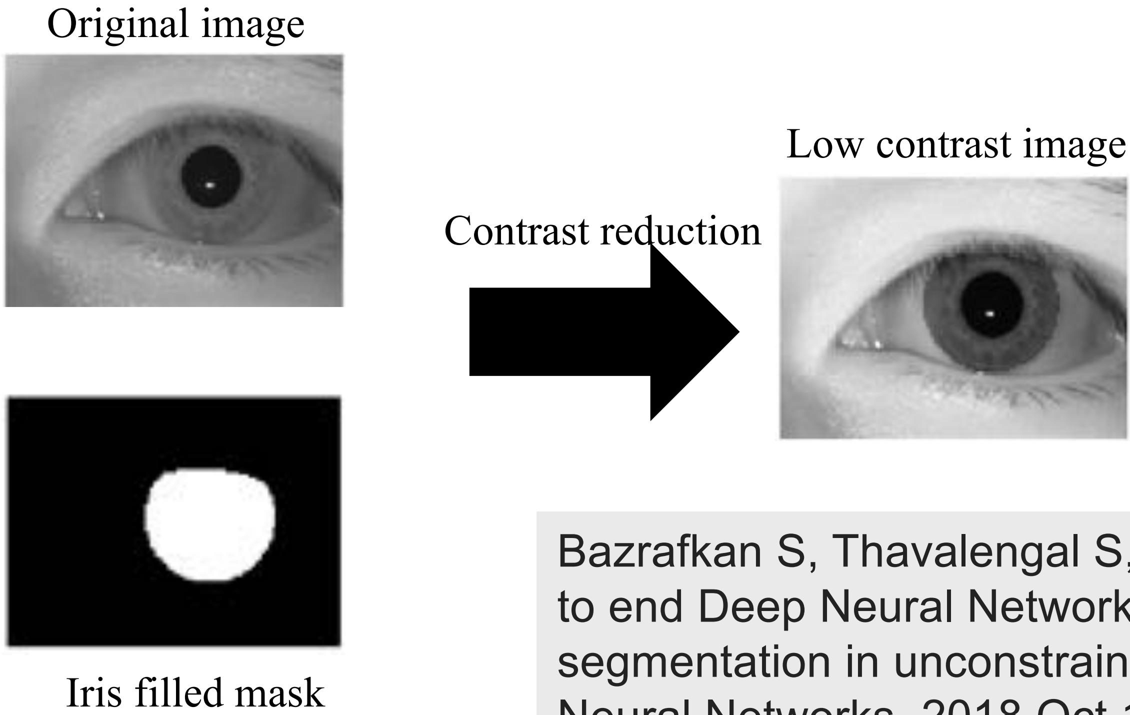
UBIRIS2

MobBio

AUGMENTATION ON TRAINING DATA – #1 - SIMPLE AUGMENTATION

CONTRAST REDUCTION

- simple augmentation – contrast reduction

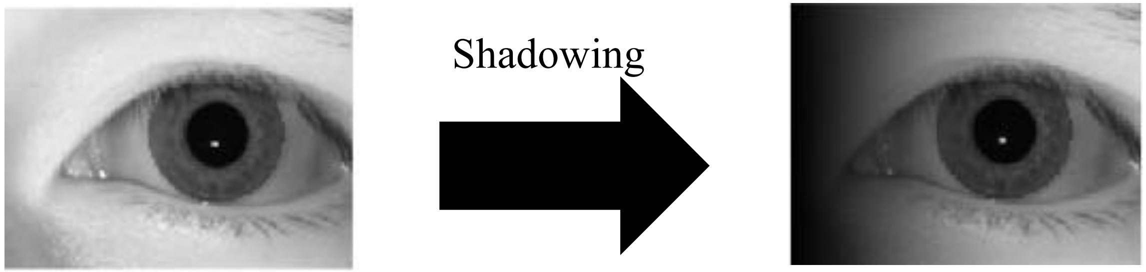


Bazrafkan S, Thavalengal S, Corcoran P. An end to end Deep Neural Network for iris segmentation in unconstrained scenarios. Neural Networks. 2018 Oct 1;106:79-95.

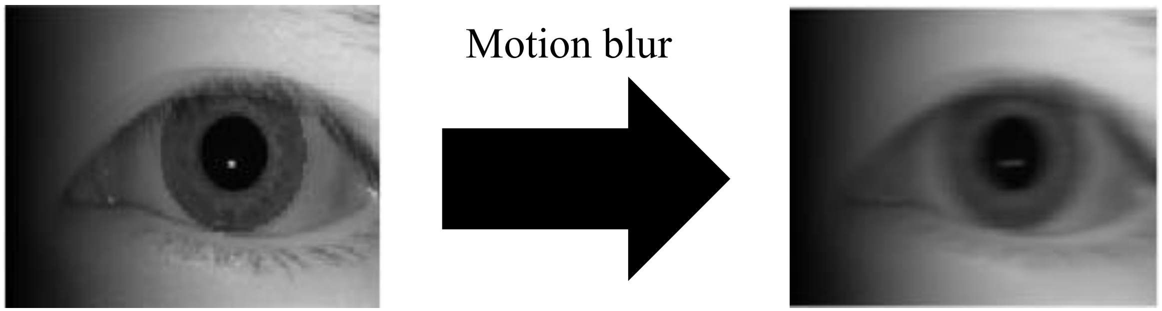
AUGMENTATION ON TRAINING DATA – #2 - COMPLEX AUGMENTATIONS

SHADOWING & MOTION BLUR

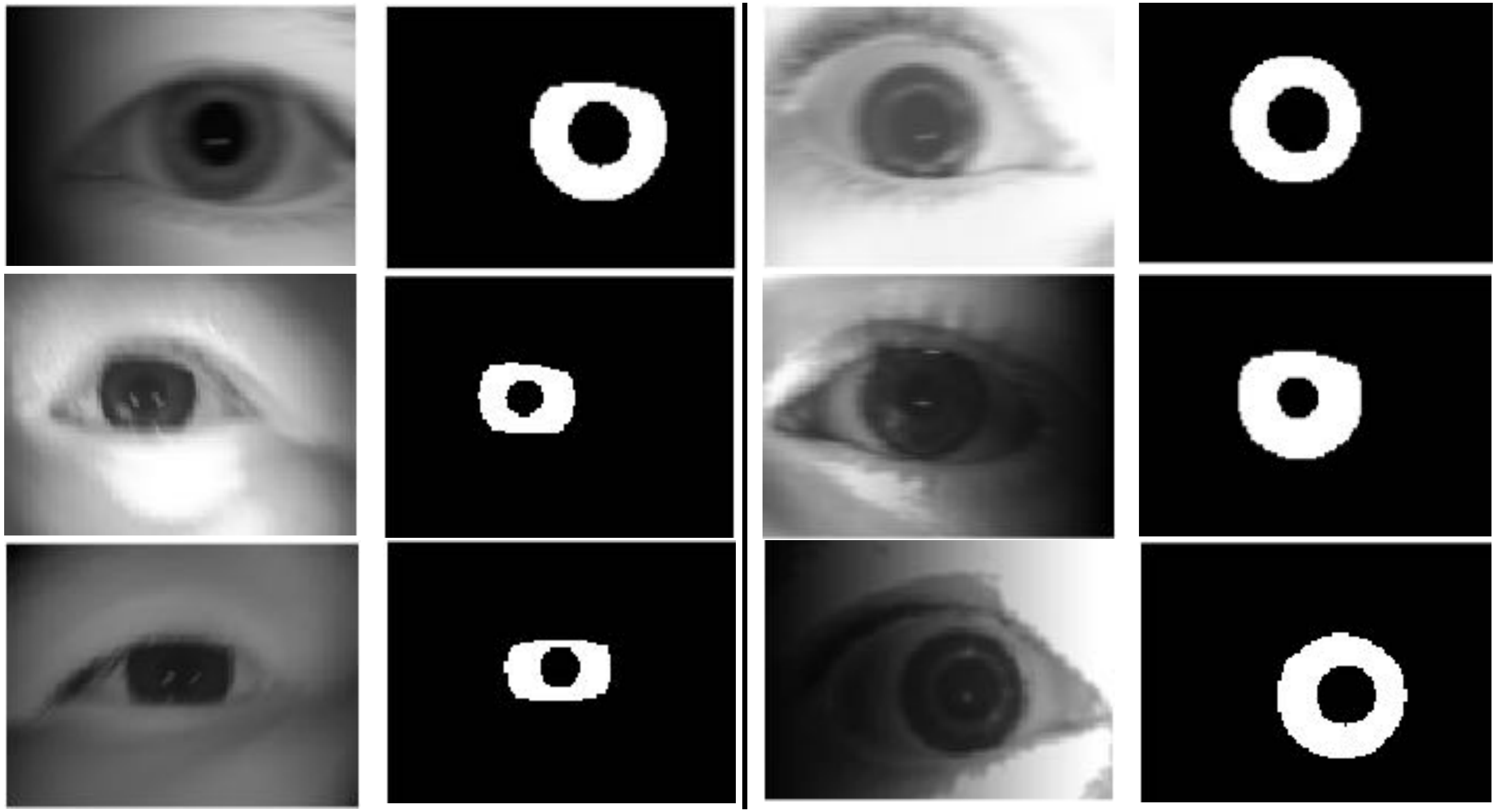
- Shadowing



- Motion Blur



AUGMENTATION ON TRAINING DATA – #3 MIXED EXAMPLES



THE 3RD **MAIN POINT** FOR TODAY!

- There are many different ways to augment data, and ….

- …. the choice of Augmentation is specific to the problem at hand (and may not be immediately obvious) …

TOPIC #2.2 SMART AUGMENTATION

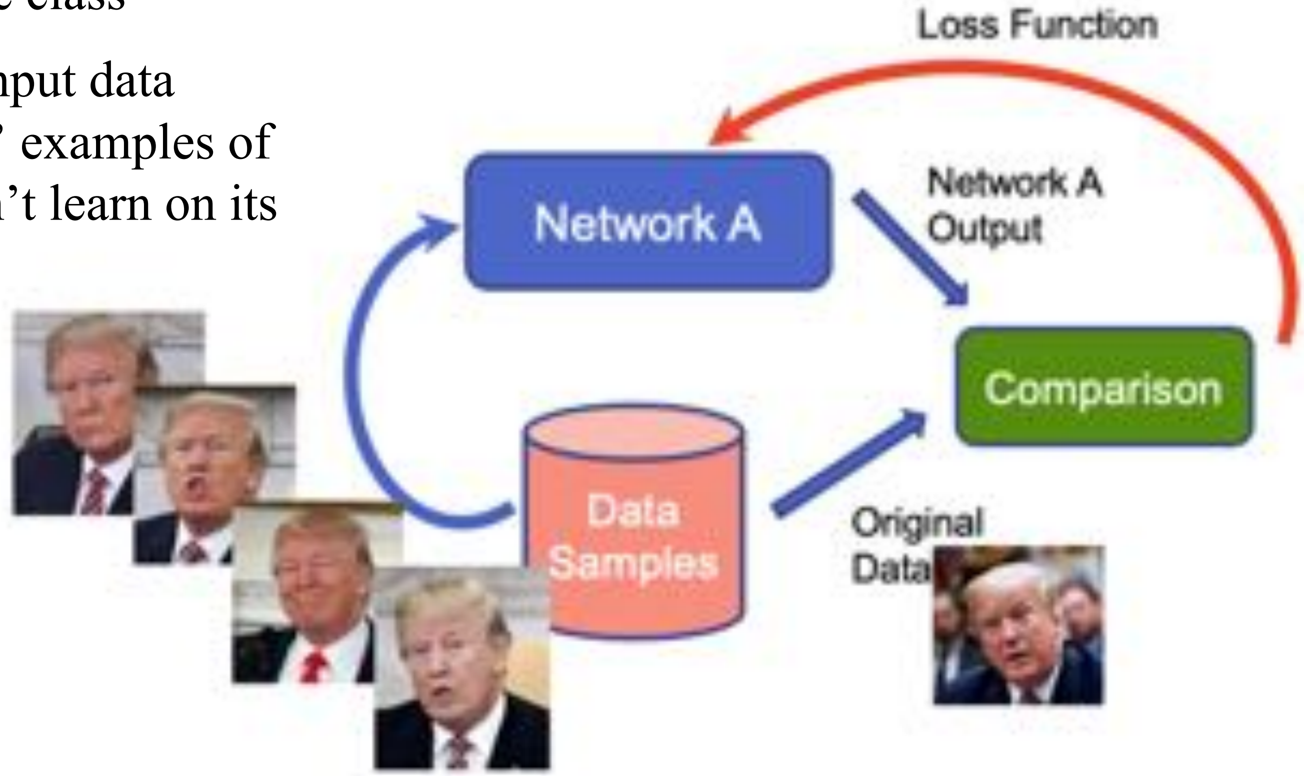
- Topic #1 – The Connection between Deep Learning & Computer Vision
  - Some Background on how I fell into the Deep Learning Pool
- **Topic #2 – The Importance of Data for Deep Learning**
  - What is Data Augmentation & why we need it
  - **Smart-Augmentation – learning new ways to augment datasets**
  - Boundary Equilibrius GAN (BEGAN) – a GAN to generate Face Images
  - Latent Spaces, learning Annotations & generating Custom Data
- Topic #3 – Virtual Reality as a source of Datasets
  - Real Data is difficult …
  - Fake data is better … and the tools are here today!

CAN WE TEACH A DEEP NETWORK TO LEARN AN AUGMENTATION STRATEGY?

- Augmenting data is a very time consuming process. Selecting the best augmentation strategy is sometimes a matter of luck, and expertise.
- **Question:** *Can we do better? Can Artificial neural networks learn the augmentation task during training?*
- **Answer:** *Yes!*

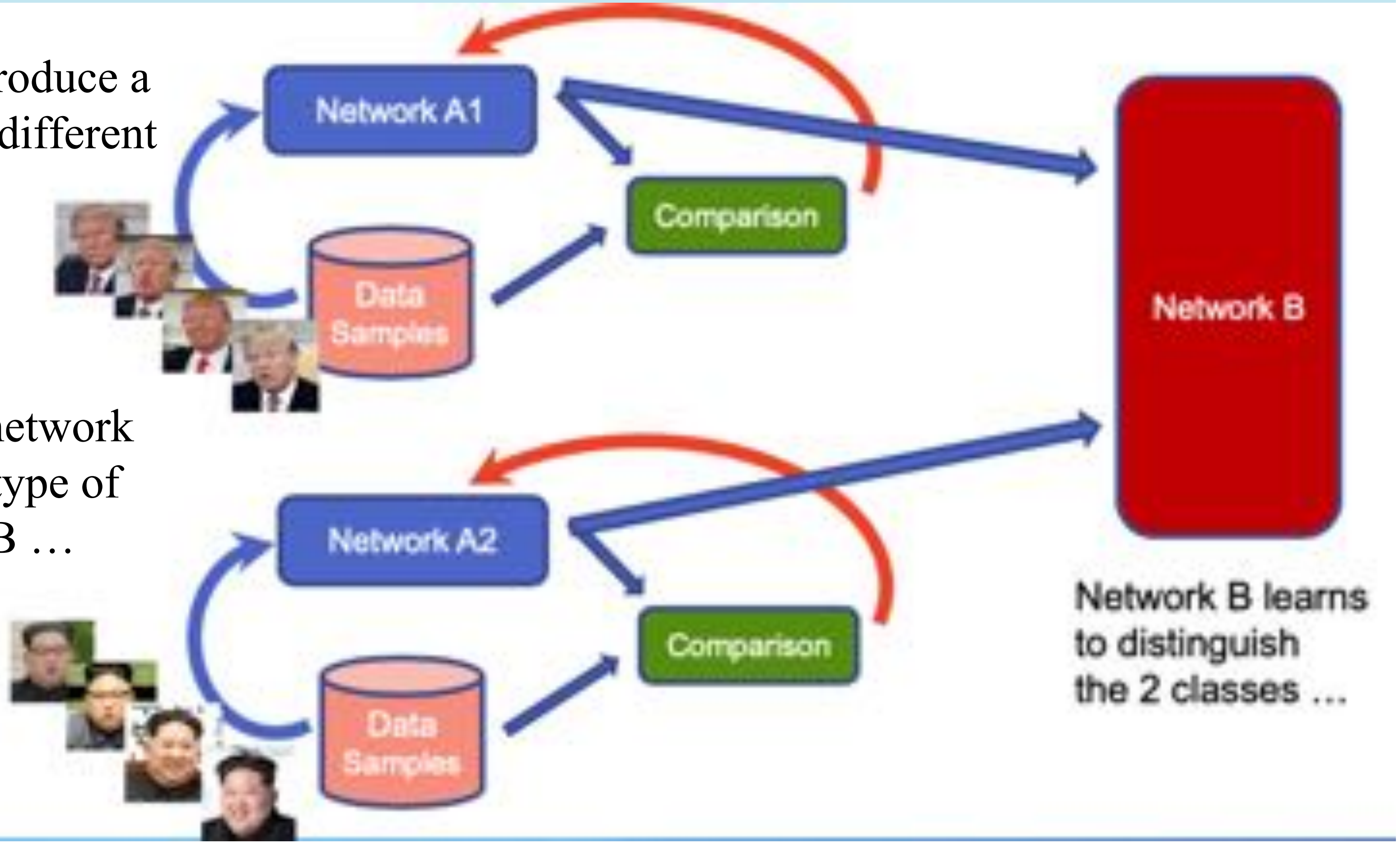
NETWORK A – LEARNING NEW DATA OF A CERTAIN CLASS

- We would like Net A to learn how to make new data of the same class
- The challenge is that the input data provides 'close to optimal' examples of class A so the network can't learn on its own ….



NOW ADD A 2ND DATA CLASS + CLASSIFIER NETWORK

- Now suppose we introduce a 2nd Network A for a different Class …
- Let us feed the two network A outputs into a 2nd type of network – Network B …

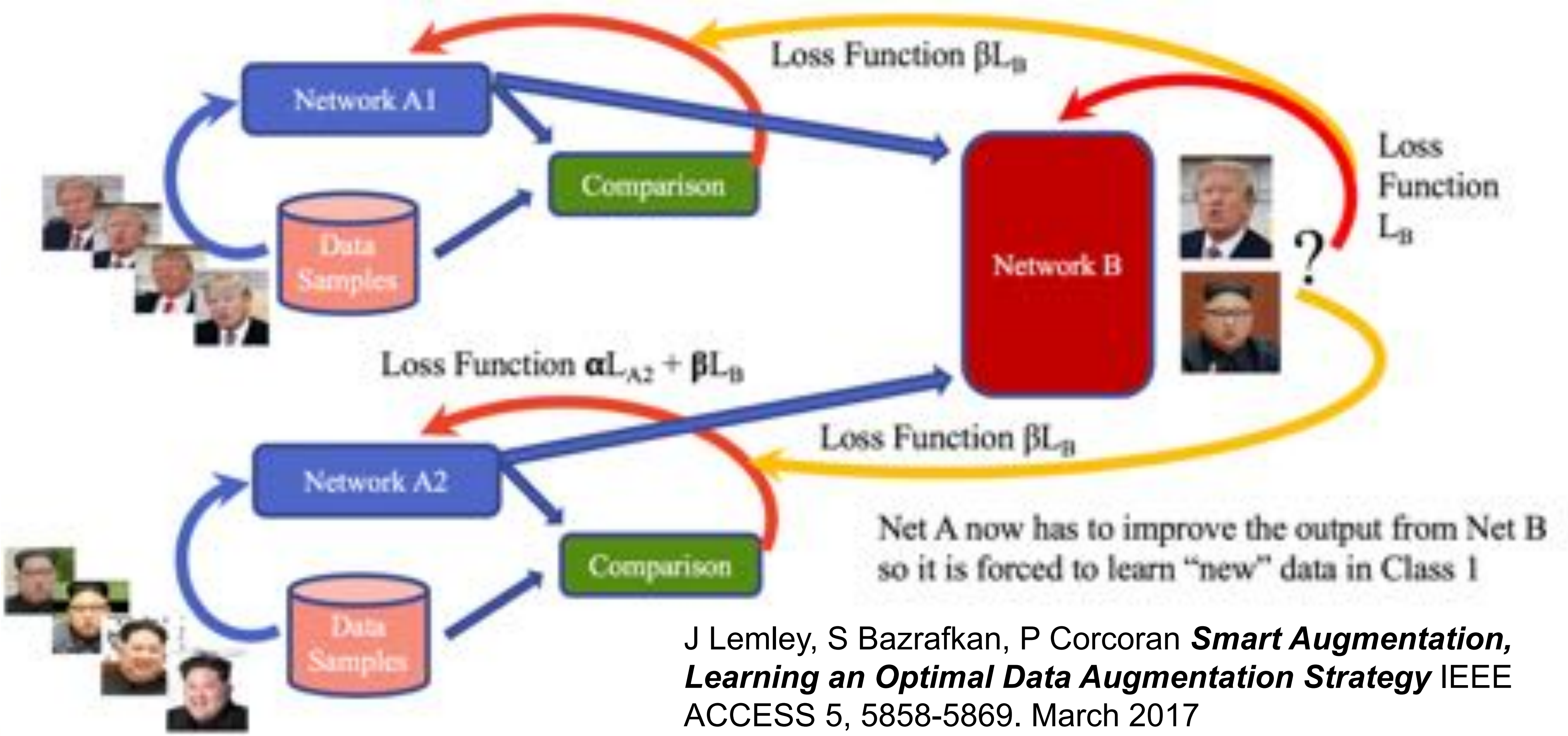


HOW DOES SMART AUGMENTATION WORK WITH IMAGES? (2)

- Network A learns to merge two or more samples in one class (in a nonlinear way) via its convolutional layers …
  - The merged sample is then used to train a target Network B
  - The loss of the target network B is fed back to inform the augmenter.
  - This generates more images for use by the target network.
- Network A (augmentor) learns to generate images that not only belong to the same class, but that improve the performance of Network B (classifier)
  - The convolutional nature of the merging of samples in Network A generates some surprising results as we'll see later …

THE FULL PICTURE …

EACH NETWORK A GETS BETTER AT GENERATING SAMPLES THAT IMPROVE NETWORK B ACCURACY



J Lemley, S Bazrafkan, P Corcoran ***Smart Augmentation, Learning an Optimal Data Augmentation Strategy*** IEEE ACCESS 5, 5858-5869. March 2017

WHAT KINDS OF IMAGES DOES NETWORK-A GENERATE?

- The images in the red box are created by a learned combination of the previous two images in that row.
- Smart Augmentation has been experimentally validated with 30 experiments which show that **Smart Augmentation:**
  - Decreases **overfitting**
  - Increases **accuracy**
  - Increases **generalization capability**
  - Can significantly *reduce the number of parameters* required to perform the same task (i.e. a *smaller network* can work just as well as a large network)

SMART AUGMENTATION WAS EFFECTIVE ON A VARIETY OF UNCONSTRAINED AND CONSTRAINED IMAGES …

AR Faces
Highly constrained

FERET
Constrained

Audience
Unconstrained

MIT Places
Highly Unconstrained

KEY RESULTS – SIGNIFICANT IMPROVEMENTS WITH SMART AUGMENTATION

A significant improvement in accuracy resulted from using Smart Augmentation.

(Green cells are Smart-Augmentation (SA) experiments; Orange cells are equivalent methods without SA)

+3.4% up to **+6.7%** on *AR Faces*

+6.1% on *Audiance*

+5.0% on *FERET dataset*

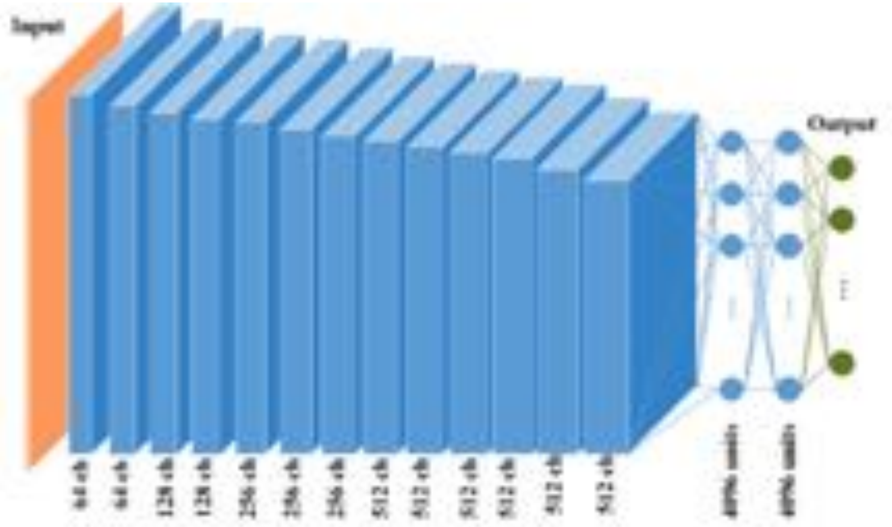
Dataset	# Net As	Input Channels	Augmented	Test Accuracy
AR Faces	1	2	no	92.50%
AR Faces	1	3	no	95.10%
AR Faces	1	4	no	91.60%
AR Faces	2	2	no	95.70%
AR Faces	2	3	no	94.20%
AR Faces	2	4	no	94.20%
AR Faces	0	NA	no	88.20%
AR Faces	0	NA	yes	89.00%
AR Faces	1	2	yes	95.70%
AR Faces	2	2	yes	95.70%
Adience	0	NA	no	70.00%
Adience	1	2	no	76.10%
FERET	0	NA	no	83.50%
FERET	1	2	no	88.50%

Citation: Joseph Lemley, Shabab Bazrafkan, and Peter Corcoran. "Smart Augmentation-Learning an Optimal Data Augmentation Strategy." *IEEE Access* (2017).

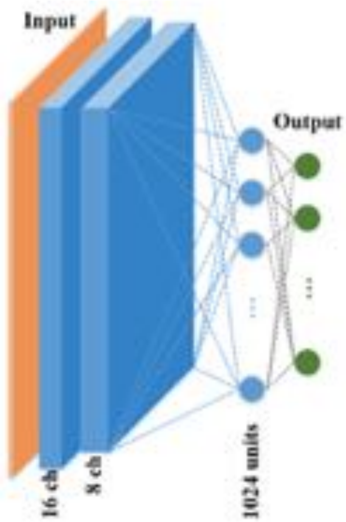
NETWORK SIZE REDUCTION VIA SMART AUGMENTATION

When trained on 2 classes from MIT Places Dataset.

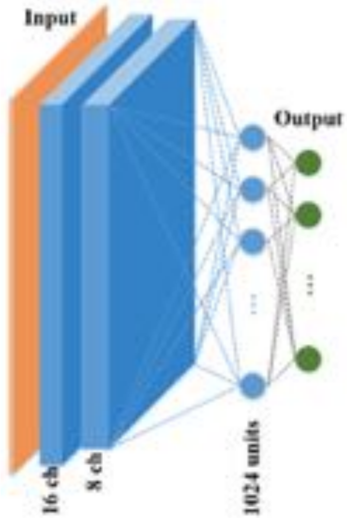
VGG 16 without SA (98.5%)



Small Network
without SA (96.5%)



Small Network
with SA (99%)



THE **4TH MAIN POINT** FOR TODAY!

- We can design neworks to improve our training data using **Smart Augmentation**, and ….
- …. they should be able to find better and more subtle augmentations than humans …

TOPIC #2.3 GENERATIVE NETWORKS

- Topic #1 – The Connection between Deep Learning & Computer Vision
  - Some Background on how I fell into the Deep Learning Pool
- **Topic #2 – The Importance of Data for Deep Learning**
  - What is Data Augmentation & why we need it
  - Smart-Augmentation – learning new ways to augment datasets
  - **Boundary Equilibrius GAN (BEGAN) – a GAN to generate Face Images**
  - Latent Spaces, learning Annotations & generating Custom Data
- Topic #3 – Virtual Reality as a source of Datasets
  - Real Data is difficult …
  - Fake data is better … and the tools are here today!

TRAINING ADVERSARIAL SAMPLES

(GOODFELLOW 2014) …

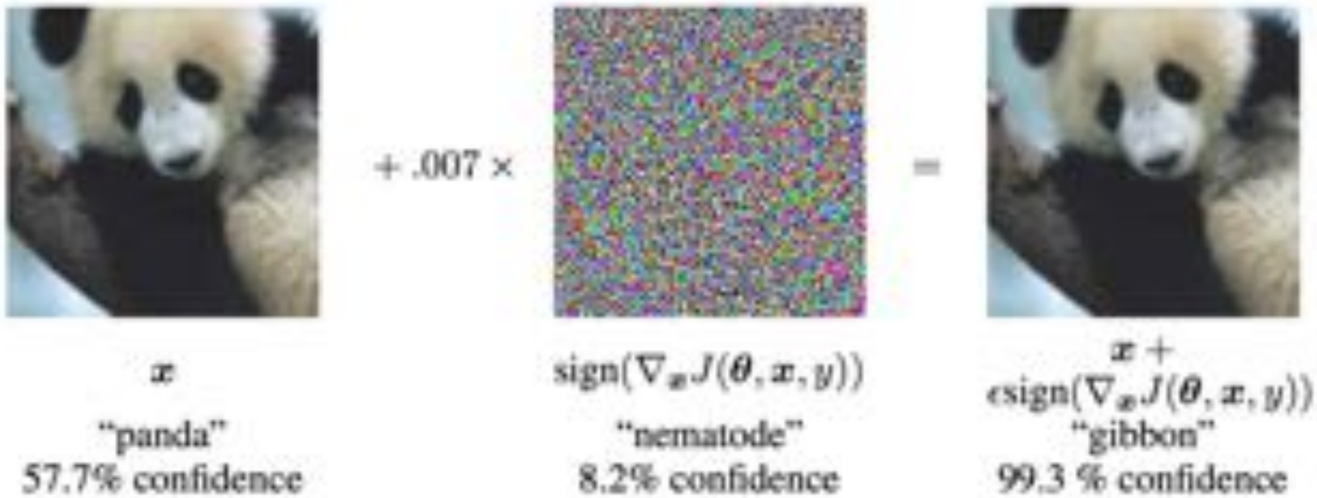
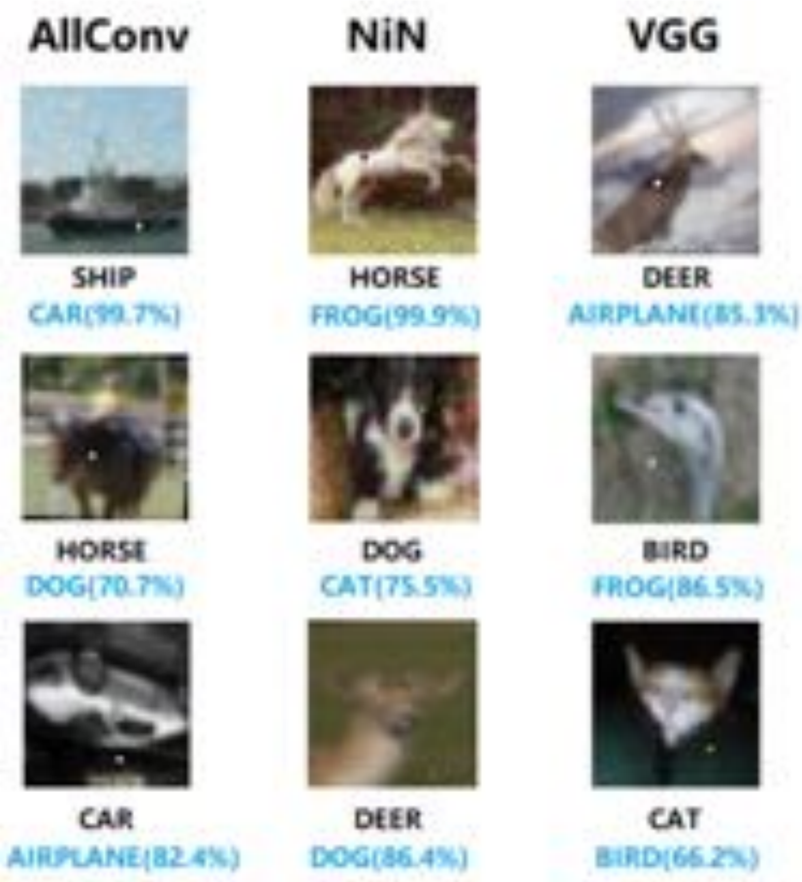


Figure 1: A demonstration of fast adversarial example generation applied to GoogLeNet (Szegedy et al., 2014a) on ImageNet. By adding an imperceptibly small vector whose elements are equal to the sign of the elements of the gradient of the cost function with respect to the input, we can change GoogLeNet's classification of the image. Here our ϵ of .007 corresponds to the magnitude of the smallest bit of an 8 bit image encoding after GoogLeNet's conversion to real numbers.

Goodfellow IJ, Shlens J, Szegedy C. Explaining & harnessing adversarial examples. arXiv:1412.6572. 2014 Dec 20.

SINGLE PIXEL ADVERSARIAL EXAMPLES …

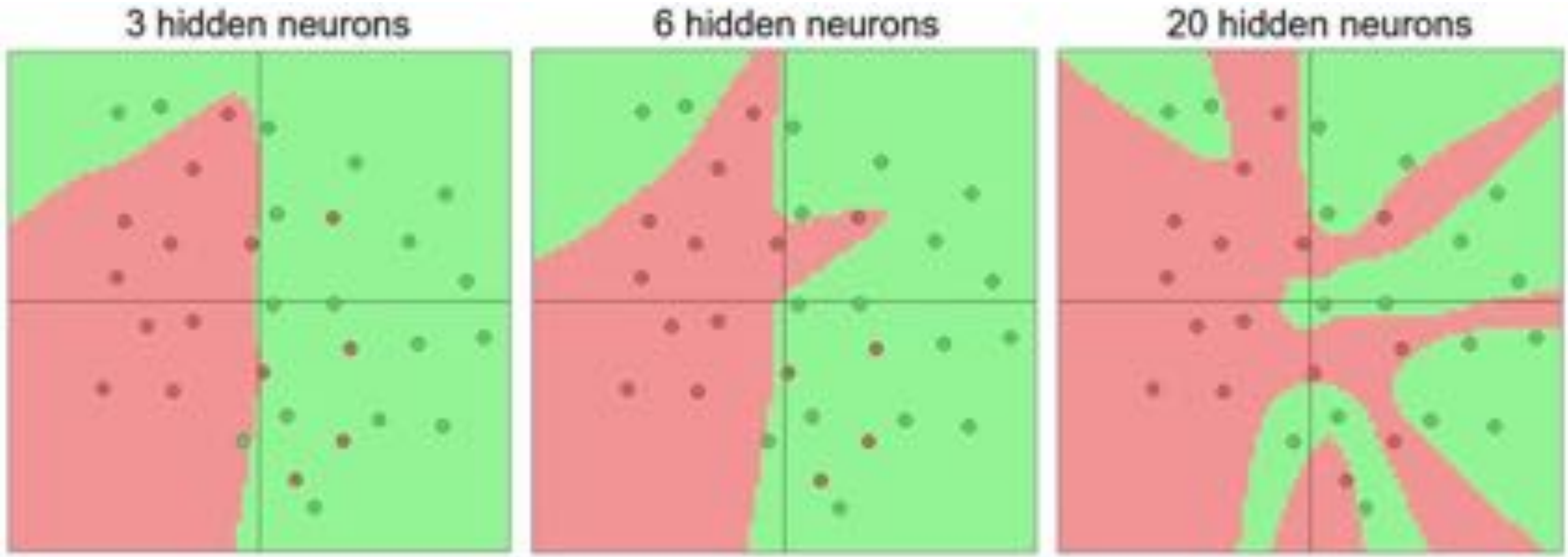


- There are many ways to ‘fool’ DNNs
- These authors showed that a single pixel in the right place can ‘trigger’ a DNN to generate false classifications
- They showed that several well-know object classification frameworks were vulnerable to this simple ‘attack’ …

Su J, Vargas DV, Kouichi S. One pixel attack for fooling deep neural networks. arXiv:1710.08864. 2017 Oct 24.

HOW DOES ADVERSARIAL TRAINING WORK …

- Adversarial networks are trained to find anomalies on the periphery of the valid data space … outliers that are wrong, but manage to trigger the network …

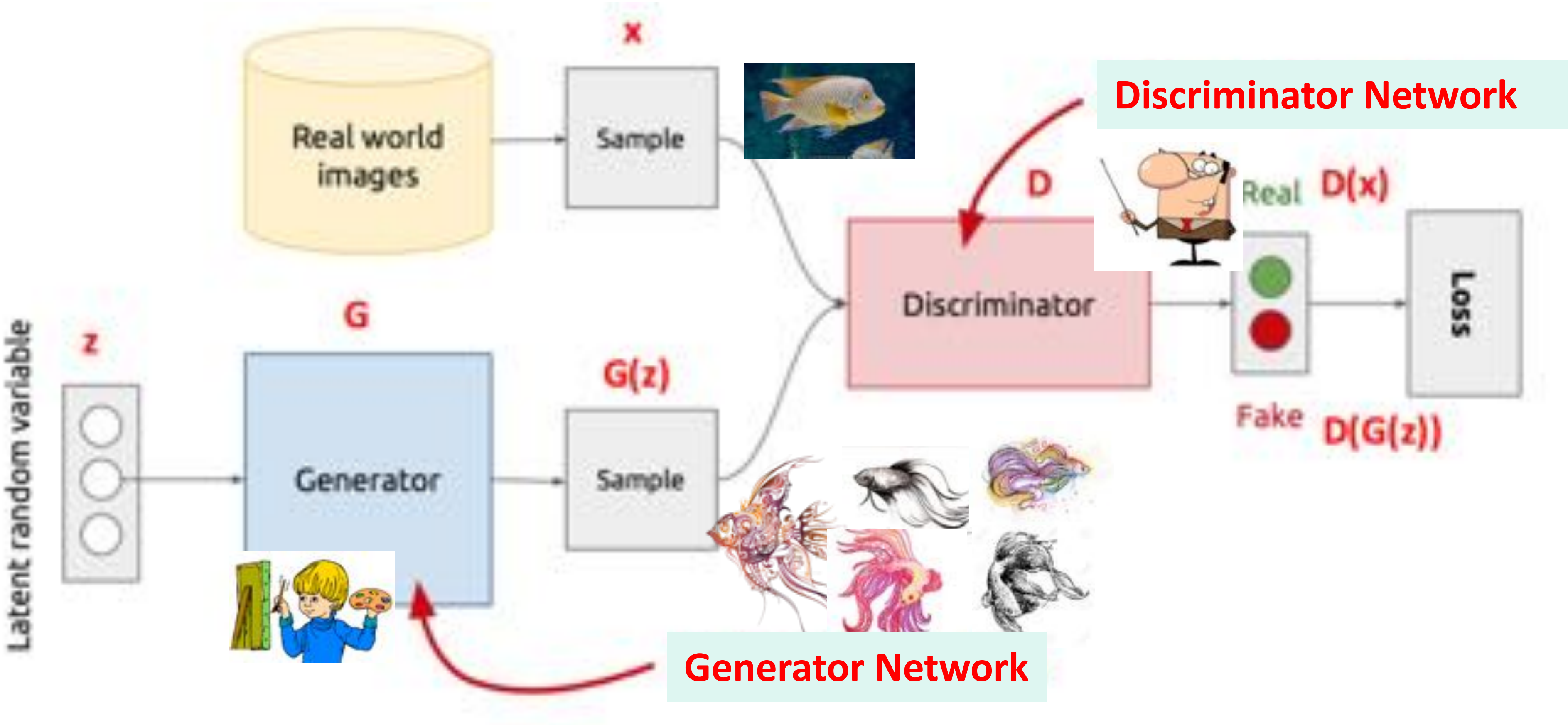


GENERATIVE ADVERSARIAL NETWORK (*GAN*) – USES ADVERSARIAL DATA TO IMPROVE THE DATA MODEL!

- Generated '*adversarial*' samples fool a discriminative model …
- **But** these *adversarial* samples can be utilized in the training process as ***false positive samples*** to make the discriminative model more robust!
  - DNN network learns the many difference between adversarial & genuine data …
  - More effort is **then** required by the *adversial generator* to determine *better adversarial samples* ...
  - Iterating this process yields an *improved* ***discriminative model*** …
    - …. and an *improved* ***generator*** for adversarial samples!

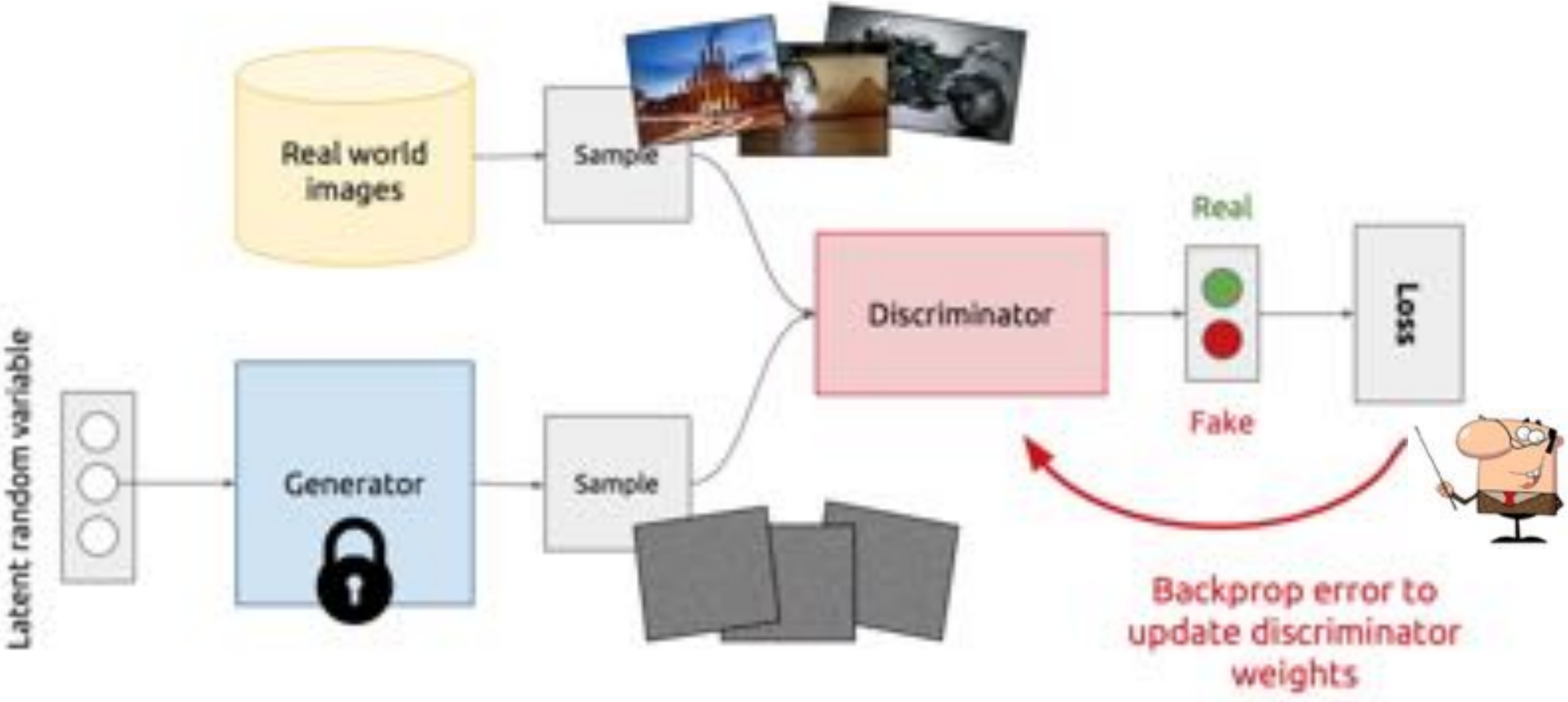
ARCHITECTURE OF GAN

#1 - *LEARNING TO GENERATE SAMPLES*



ARCHITECTURE OF GAN

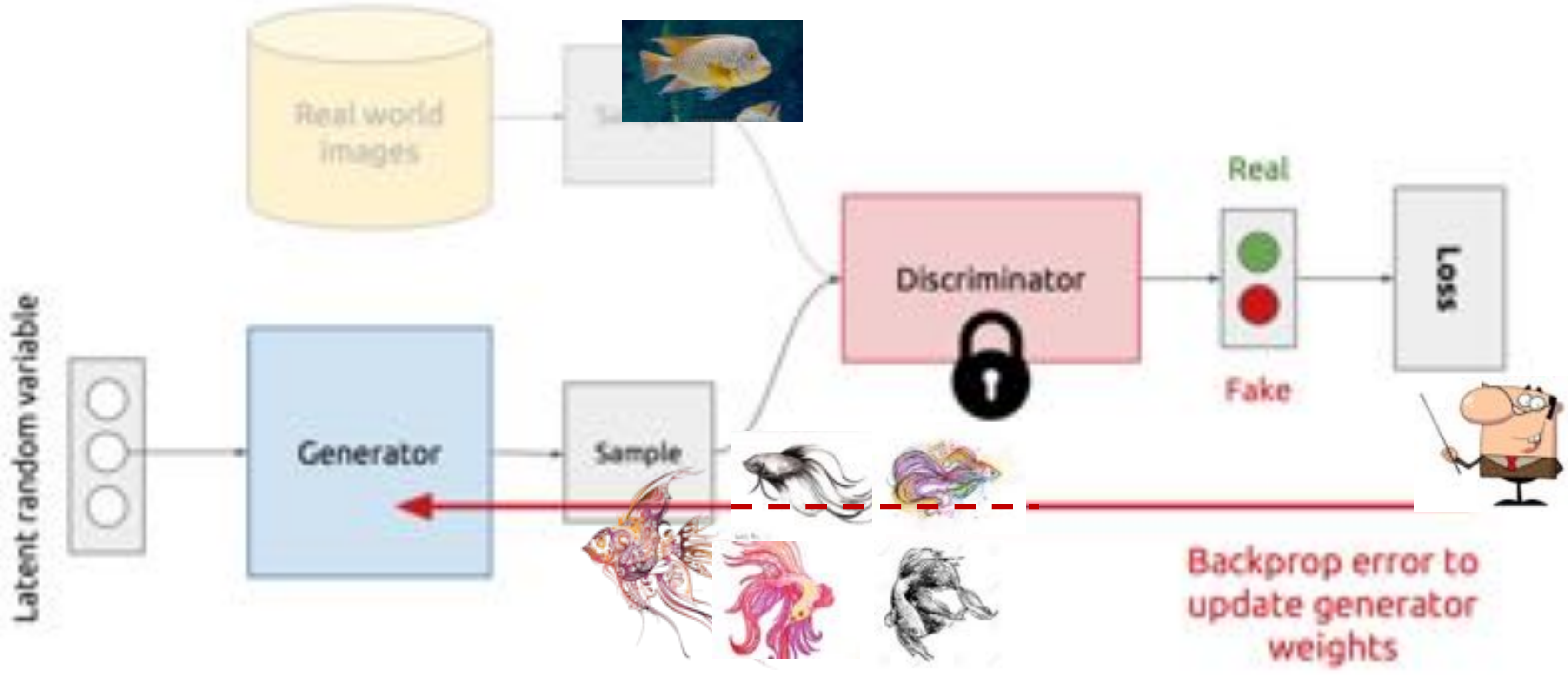
#2 - SEPARATE TRAINING NEEDED FOR THE DISCRIMINATOR …



https://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-generative-models-and-adversarial-training-upc-2016

ARCHITECTURE OF GAN

#3 … AND THE GENERATOR MODULES!



https://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-generative-models-and-adversarial-training-upc-2016

GAN IS FORMULATED AS A MIN/MAX GAME …

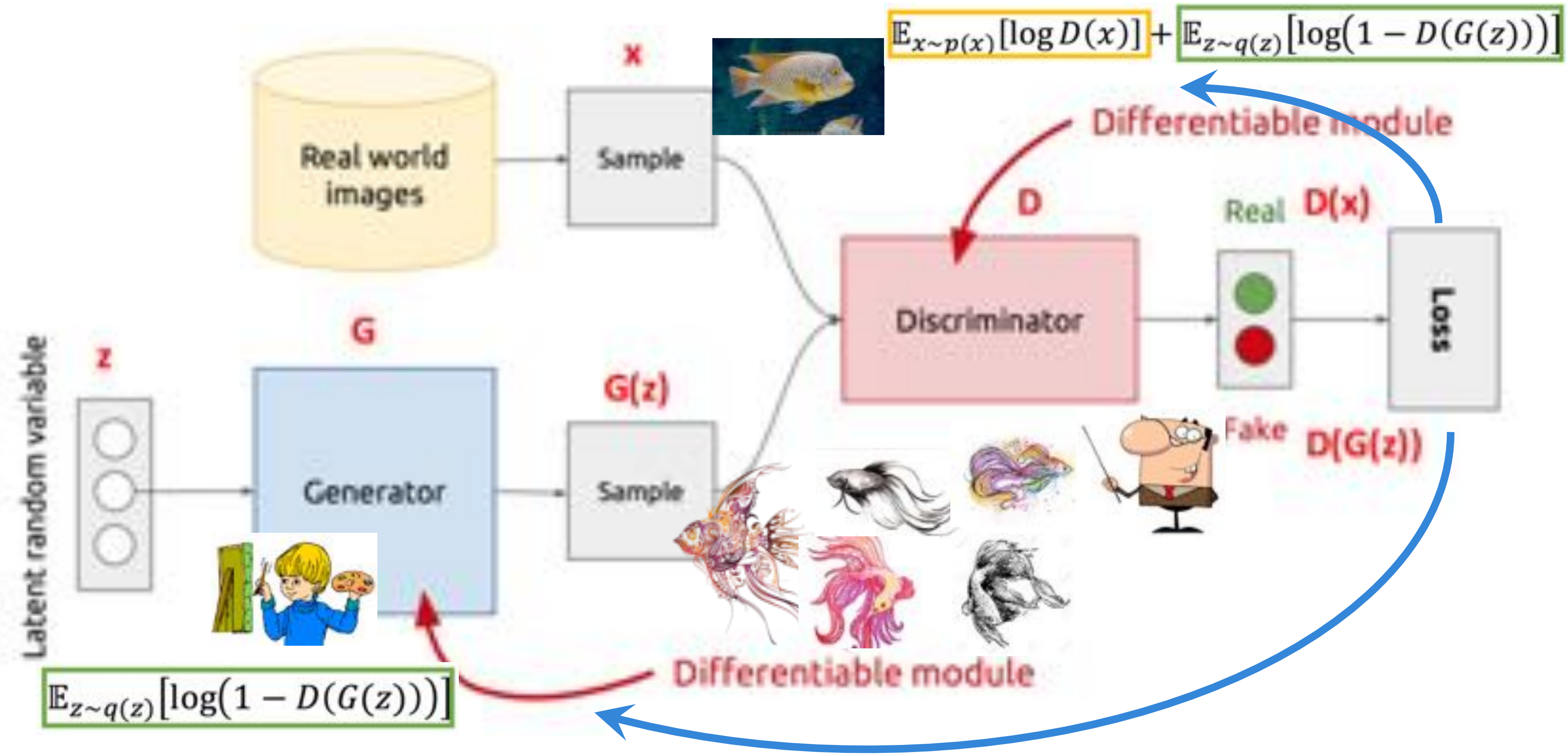
GAN's formulation

$$\min_G \max_D V(D, G)$$

- It is formulated as a **minimax game**, where:
  - The Discriminator is trying to maximize its reward $\boldsymbol{V(D, G)}$
  - The Generator is trying to minimize Discriminator's reward (or maximize its loss)

$$V(D, G) = \boxed{\mathbb{E}_{x \sim p(x)}[\log D(x)]} + \boxed{\mathbb{E}_{z \sim q(z)}\left[\log\left(1 - D(G(z))\right)\right]}$$

FULL ARCHITECTURE OF GAN …



THE 5.1TH **MAIN POINT** FOR TODAY!

- Generative Adversarial Networks (GANs) are very powerful tools for training data generators ….

- …. Lets take a look at how we can use them to create Face Data …

CELEBFACES ATTRIBUTES (CELEB-A)

SoTA : A large-scale face attributes dataset; 200K+ celebrity images; 40 attribute annotations per image; images cover large pose variations and background clutter.

CelebA has large diversities, large quantities, and rich annotations, including:

10,177 number of identities,

202,599 number of face images,

5 landmark locations, 40 binary attributes annotations per image

The dataset can be employed as the training and test sets for the following computer vision tasks: face attribute recognition, face detection, and landmark (or facial part) localization.

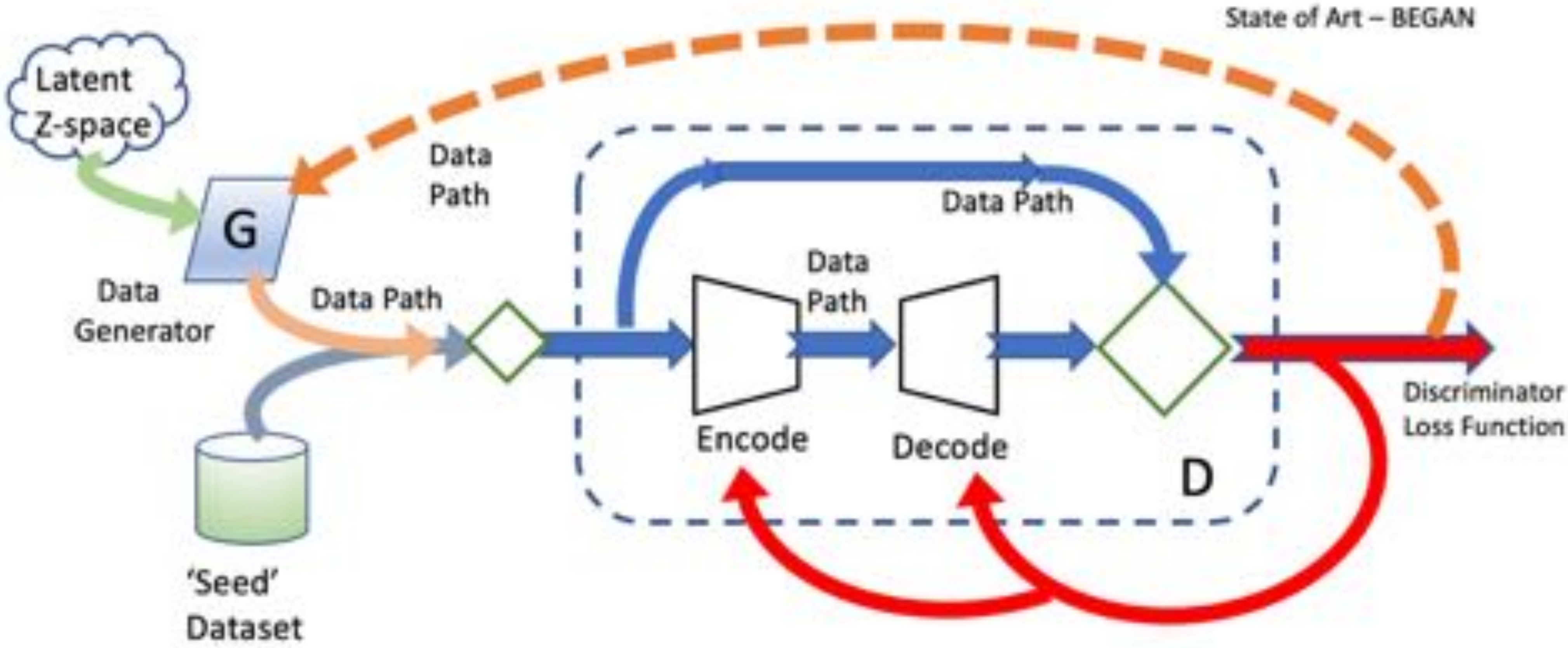
More Info & Download Utilities:
http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html

Size: > 1 GB (compressed)
Number of Records: 202,599 facial images with annotations

Sample Images

BOUNDARY EQUILIBRIUM GAN *(BEGAN)*

*REPLACES CLASSIC DISCRIMINATOR WITH AN **AUTO-ENCODER** …*



THE LOSS FUNCTIONS HAVE COMPLEX RELATIONSHIPS …

https://blog.heuritech.com/2017/04/11/began-state-of-the-art-generation-of-faces-with-generative-adversarial-networks/

Here is the complete BEGAN objective:

$$\mathcal{L}_D = \mathcal{L}(x) - k_t.\mathcal{L}(G(z))$$

[Max/Min goal - +ve reconstruction for real images; -ve for generated]

$$\mathcal{L}_G = \mathcal{L}(G(z))$$

$$k_{t+1} = k_t + \lambda * (\gamma.\mathcal{L}(x) - \mathcal{L}(G(z)))$$

[balancing of each iteration based on diversity & gain of learning rate]

- $\mathcal{L}_D$ and $\mathcal{L}_G$ are the respective losses for *D* and *G* (what they try to minimize).
- $\mathcal{L}_D$ is only used to optimize θ_D and $\mathcal{L}_G$ is only used to optimize θ_G.
- $\mathcal{L}(x)$ and $\mathcal{L}(G(z))$ are the losses of reconstruction of real and generated images.
- γ is the diversity ratio (in $[0, 1]$) defined before as: $\gamma = \mathbb{E}[\mathcal{L}(G(z))]/\mathbb{E}[\mathcal{L}(x)]$.
- k_t is the adaptive term that will allow us to balance the losses
- λ is the proportional gain for k_t (aka the learning rate for k_t).

BEGAN CAN GENERATE 128 x 128 RANDOM FACIAL CROPS FROM LARGE LATENT Z-SPACE

- Boundary Equilibrium GAN successfully generates anatomically coherent faces at a resolution of 128×128 pixels
- As of May 2017, this was the state of the art

THE 5.2TH **MAIN POINT** FOR TODAY!

- State of art GAN methods can generate very photo-realistic face samples …

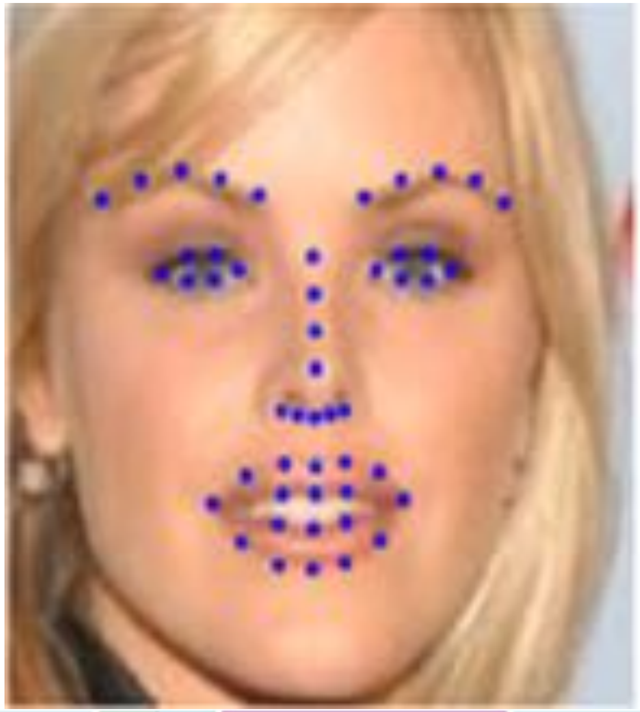
- …. although it did take a clever PhD student 3+ months to understand & learn how to tweak everything!

TOPIC #2.4 GENERATING ANNOTATIONS WITH LATENT SPACES

- Topic #1 – The Connection between Deep Learning & Computer Vision
  - Some Background on how I fell into the Deep Learning Pool
- **Topic #2 – The Importance of Data for Deep Learning**
  - What is Data Augmentation & why we need it
  - Smart-Augmentation – learning new ways to augment datasets
  - Boundary Equilibrius GAN (BEGAN) – a GAN to generate Face Images
  - **Latent Spaces, learning Annotations & generating Custom Data**
- Topic #3 – Virtual Reality as a source of Datasets
  - Real Data is difficult …
  - Fake data is better … and the tools are here today!

THE CHALLENGE WE POSED …

- Can we modify BEGAN to generate random faces that are accurately "annotated" … ?
  - Why would you want to do that?



- It becomes possible to generate **random faces** for **training data** …
- Can '**prove**' they are not the 'original' faces – solves **Privacy** issues!
- In theory can generate **more faces** than in the original dataset … (???)

DATASET AND INITIAL LANDMARK DETECTION

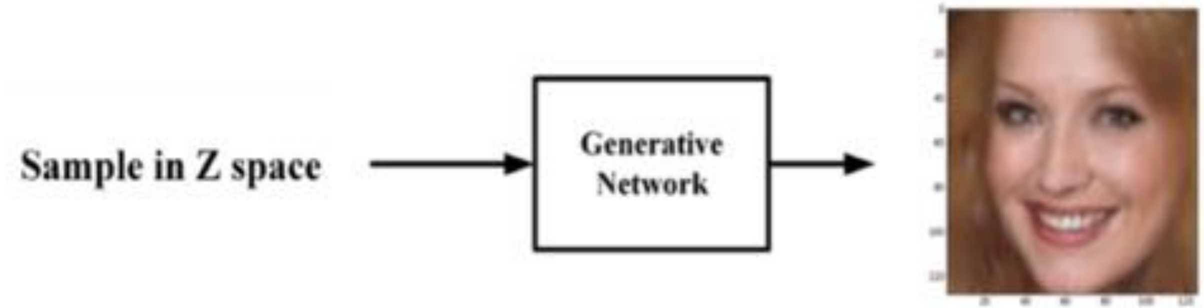
- The **CelebA dataset - 202,599** original images with 40 unique attributes is used for training our GAN framework *(http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html)*
- **OpenCV** frontal face cascade classifier detects facial regions
  - cropped and resized to 128×128 pixels
- Initial landmark detection is performed using the method developed by Astana *et. al* due to its ability to be effective on unconstrained faces
  - Landmark detector estimates a set of **49 landmark points** defined on the contours of eyebrows, eyes, mouth and the nose

Asthana A, Zafeiriou S, Cheng S, Pantic M. ***Incremental face alignment in the wild****. In proceedings of the IEEE conference on computer vision and pattern recognition 2014 (pp. 1859-1866).*

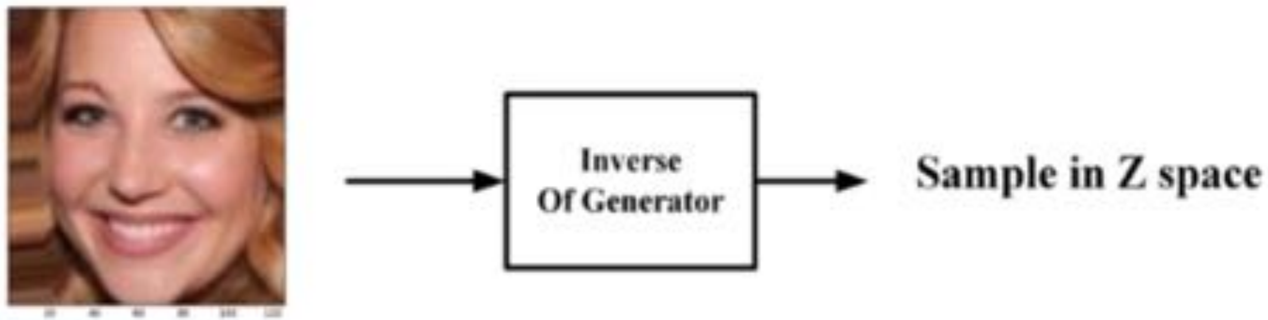
LANDMARK EXAMPLES ON THE ORIGINAL CELEB-A DATASET …

MAPING OF CELEBA DATA SAMPLES INTO Z-SPACE

- A Generator, ***G***, is trained to produce random faces using BEGAN

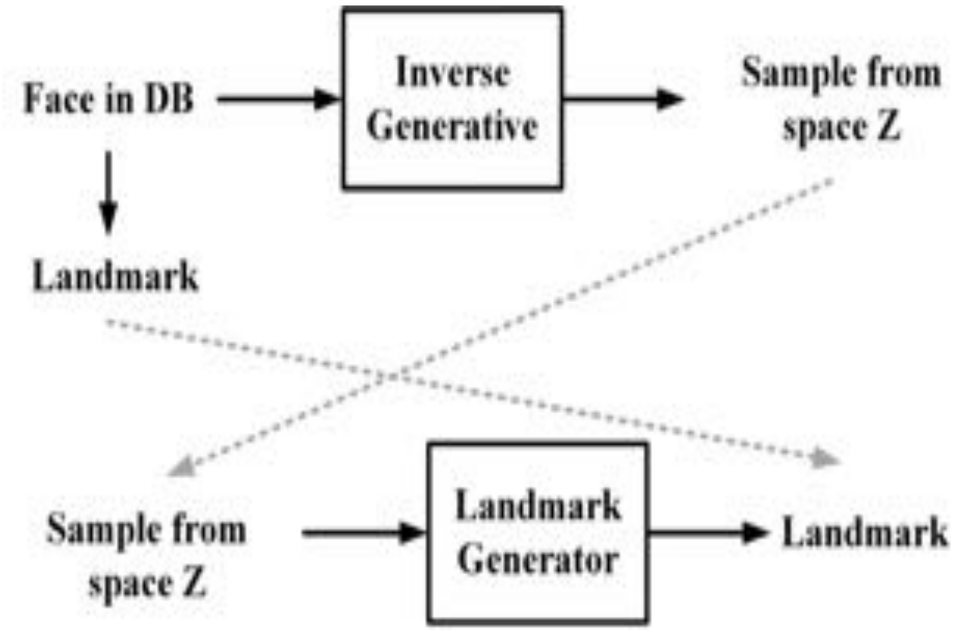


- A new *Inverse Generator* maps a face sample back into the latent *Z-space* so we can associate a set of landmarks with a 'location' in Z-space …



TRAINING THE LANDMARK GENERATOR …

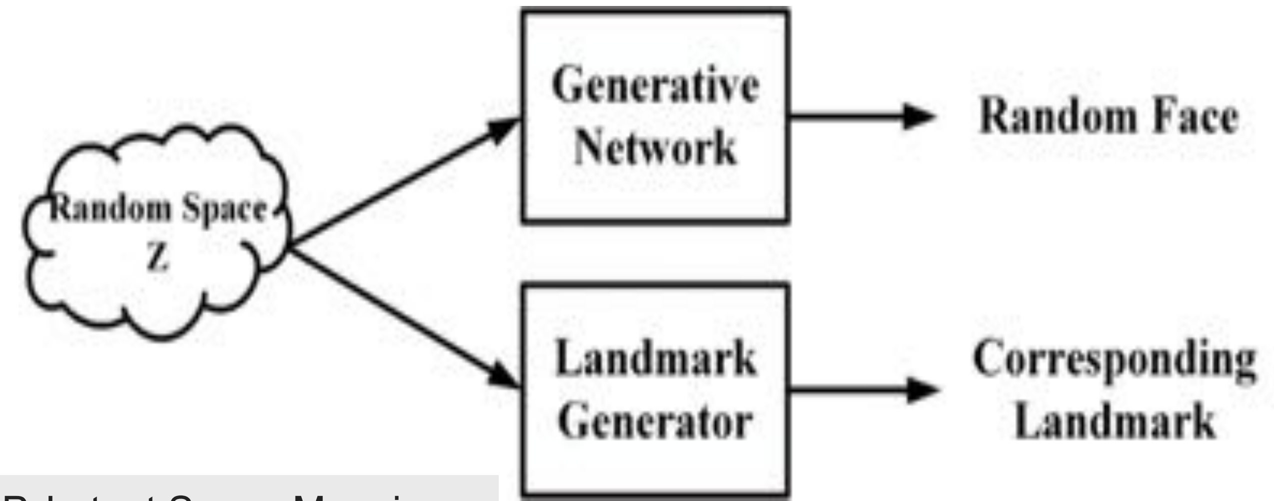
- CelebA data samples are mapped to Z-space to Train the Landmark Generator
- Landmark Generator is trained to learn annotations from latent Z-space



1. A Facial image sample is mapped into the latent space by the *Inverse Generator*
2. The corresponding facial *landmark data* becomes a training sample associated with corresponding location in Z-space; (repeat 200k times)
3. A *generator* is trained to map Z-space locations to corresponding landmark point set
4. If landmark data is accurate *ground truth* the generator can handle wide variations in illumination & pose

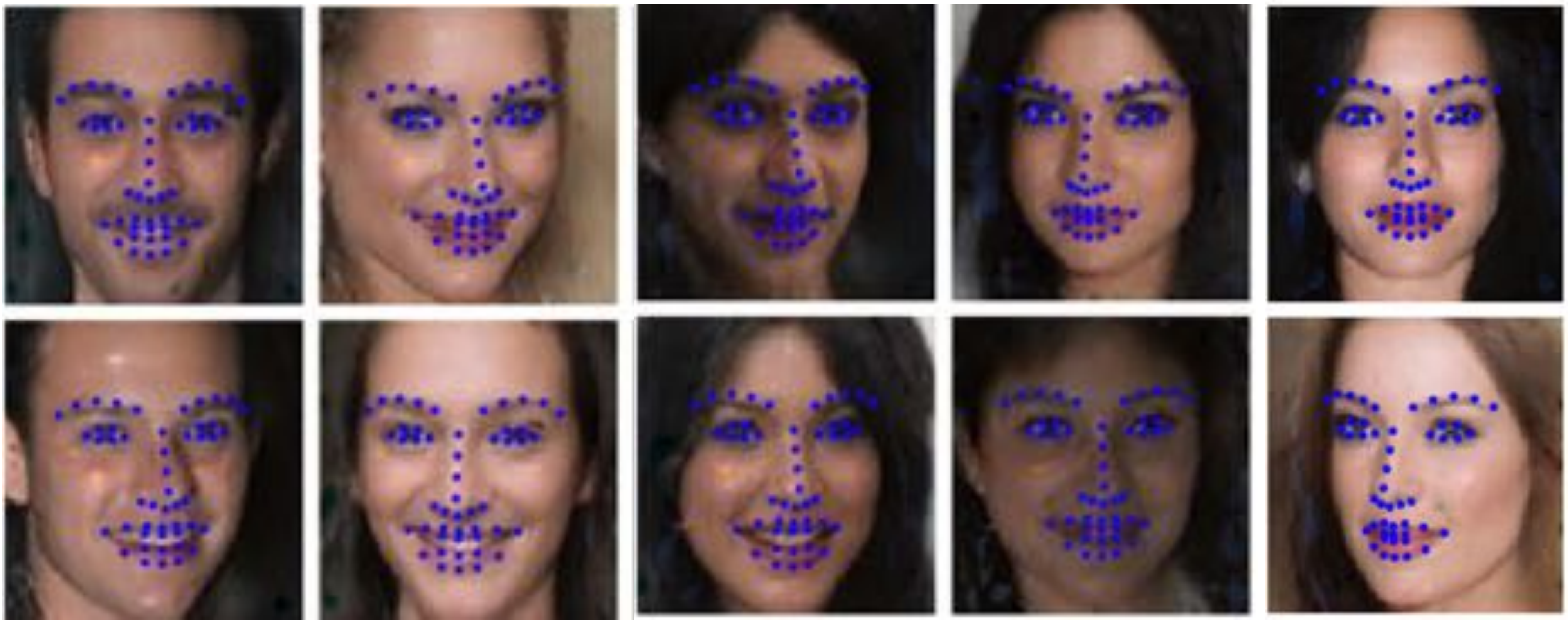
FINALLY, FACES GENERATED TOGETHER WITH ANNOTATIONS …

- Generate a random Face image from the Latent Z-space …
- Matching annotations are generated directly from the originating Z-space vector values
- First time to generate *data + annotations* from a common Z-space?

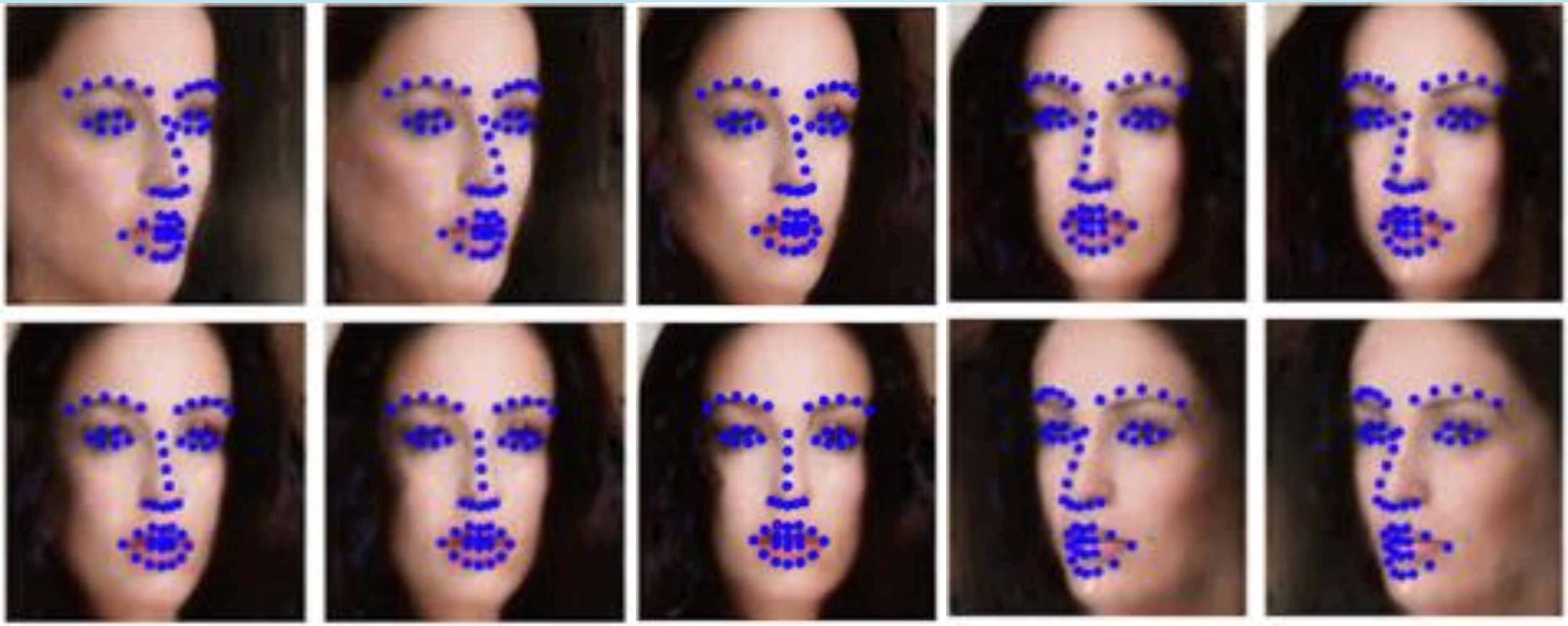


Bazrafkan S, Javidnia H, Corcoran P. Latent Space Mapping for Generation of Object Elements with Corresponding Data Annotation. Pattern Recognition Letters. 2018 Oct 23.

EXAMPLE #1 OF GENERATED DATASET MULTIPLE DIFFERING FACIAL IMAGES



EXAMPLE #2 OF GENERATED DATASET
SAME BASE FACE WITH DIFFERING FACIAL POSE



SOME OBSERVATIONS

- The learned mapping into latent Z-Space will vary according to the training dataset and the DNN structures employed in the GAN
  - The learned Z-Space characteristics depend on the training data, ***but can encapsulate pose & lighting variations*** – key challenges for face generation
  - Examples of pose variation step incrementally between 2 end-points in the Z-Space but the finer details still elude us … our understanding is still a 'work-in-progess' …
- The latent Z-space that is trained comprises 64 x 32bit floating point numbers has, potentially 2^38 discriminating capability … so, IN THEORY …
  - **274B** potential samples (c.30 times global population!) …

THE 6.1TH **MAIN POINT** FOR TODAY!

- Possible to generate random facial datasets with annotations for relatively little cost/effort ….

  https://www.youtube.com/watch?v=PWdT3Q5T5U8

- …. if only we could control how these are generated …

MORE RECENT UPDATES

BEYOND RANDOM FACES #1

VERSATILE AUXILLIARY CLASSIFIER/REGRESSION TECHNIQUES

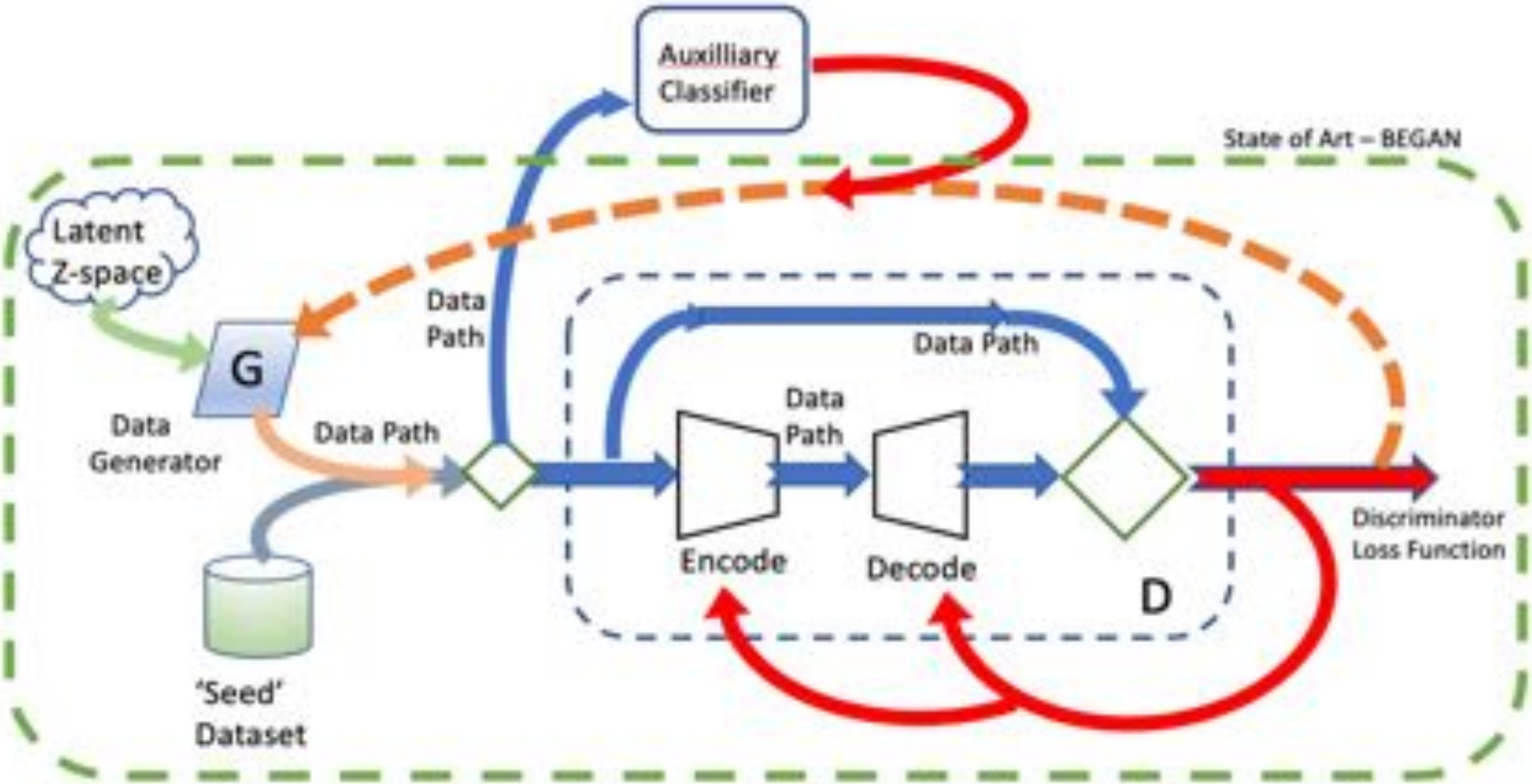
- We can now 'learn' to generate faces with particular attributes e.g. male or female faces; age categories: teen-ager, young adult, older adult, senior; etc, etc …
- By adding 'regulation' into the feedback look it is possible, for example, to generate random faces with a specific pose characteristic …. *for more info:*

#1 Bazrafkan S, Javidnia H, Corcoran P. Versatile Auxiliary Classifier+ Generative Adversarial Network (VAC+ GAN); Training Conditional Generators. arXiv preprint arXiv:1805.00316. ***2018 May 1***

#2 Bazrafkan S, Corcoran P. Versatile Auxiliary Regressor with Generative Adversarial network (VAR+ GAN). arXiv preprint arXiv:1805.10864. ***2018 May 28***

#3 Bazrafkan S, Corcoran P. Versatile Auxiliary Classifier with Generative Adversarial Network (VAC+ GAN), Multi Class Scenarios. arXiv preprint arXiv:1806.07751. ***2018 Jun 19***

BEYOND RANDOM FACES #2A
VERSATILE AUXILLIARY CLASSIFIER TECHNIQUE



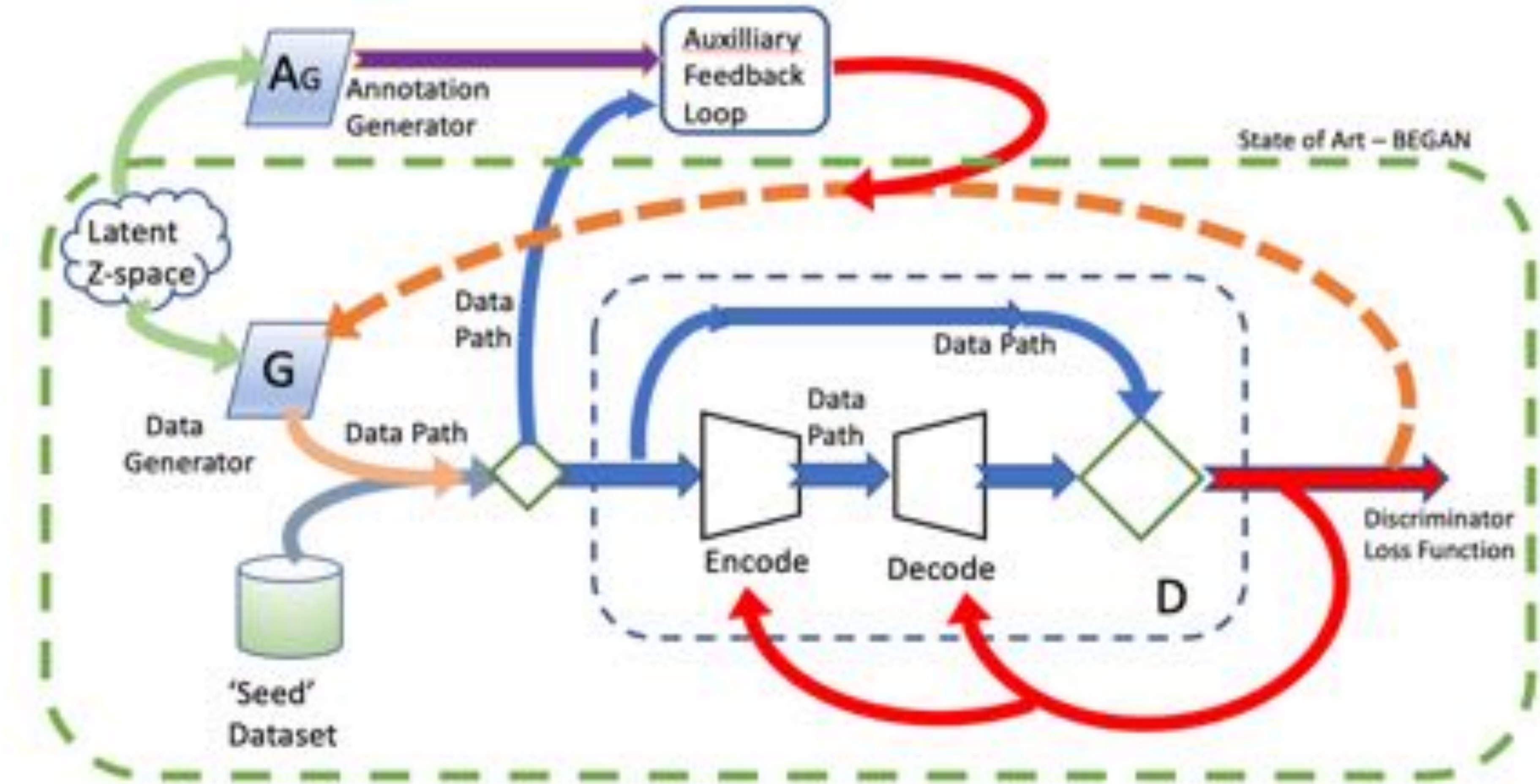
BEYOND RANDOM FACES #2B

VERSATILE AUXILLIARY CLASSIFIER TECHNIQUE

FEMALE SAMPLES

MALE SAMPLES

BEYOND RANDOM FACES #3A
VERSATILE AUXILLIARY REGRESSION TECHNIQUE



BEYOND RANDOM FACES #3B
VERSATILE AUXILLIARY REGRESSION TECHNIQUE

GENERATED DATA SAMPLES CONSTRAINED TO A SPECIFIC SET OF LANDMARK POINTS

(a) Proposed method (VAR+GAN).

(b) cBiGAN.

THE 6.2TH **MAIN POINT** FOR TODAY!

- Possible to generate facial datasets with annotations for relatively little cost/effort ….

- …. AND target a specific ‘class’ of facial data, and/or ‘regularize’ the data to a specific pose (or illumination, etc) …

TOPIC #3 SOLVING THE PROBLEMS OF "REAL DATA" …

- Topic #1 – The Connection between Deep Learning & Computer Vision
  - Some Background on how I fell into the Deep Learning Pool
- Topic #2 – The Importance of Data for Deep Learning
  - What is Data Augmentation & why we need it
  - Smart-Augmentation – learning new ways to augment datasets
  - Boundary Equilibrius GAN (BEGAN) – a GAN to generate Face Images
  - Latent Spaces, learning Annotations & generating Custom Data
- **Topic #3 – Virtual Reality as a source of Datasets**
  - **Real Data is difficult …**
  - Fake data is better … and the tools are here today!

FAKE DATA – THE FUTURE OF COMPUTER VISION & SMART CAMERAS?

Presented by Prof. Peter Corcoran at International Smart Camera Technology Workshop

NUI Galway
OÉ Gaillimh

C3Imaging Centre, College of Engineering & Informatics,
National University of Ireland Galway

Real Data is Difficult …

YOU NEED A LOT OF IT …

MS-COCO

(MICROSOFT COMMON OBJECTS IN CONTEXT)

25 GB

SoTA : [Mask R-CNN]

COCO is a large-scale dataset for object detection, segmentation and captioning. It has several key features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 330K images (>200K labeled)
- 1.5 million object instances
- 80 object categories
- 91 stuff categories
- 5 captions per image
- 250,000 people with keypoints

More Info & Download Utilities:
https://arxiv.org/pdf/1405.0312
https://github.com/cocodataset/cocoapi

Size: ~25 GB (compressed)
Number of Records: 330K images, 80 object categories, 5 captions per image, 250,000 people with key points

OPEN IMAGES DATASET 500 GB

SoTA : Resnet 101 image classification model (trained on V2 data): Model checkpoint, Checkpoint readme, Inference code.

Open Images Dataset comprises c.9 million URLs for images.

These images have been annotated with image-level labels & bounding boxes spanning thousands of classes.

The dataset contains a training set of 9,011,219 images, a validation set of 41,260 images and test set of 125,436 images.

News Flash – V4 released earlier this year has 15,440,132 boxes on 600 categories; 30,113,078 image-level labels on 19,794 categories; https://www.analyticsvidhya.com/blog/2018/05/google-released-latest-open-images-dataset/

More Info & Download Utilities:

(v 3) https://github.com/openimages/dataset
(v 4) https://storage.googleapis.com/openimages/web/index.html

Size: ~500 GB (compressed)
Number of Records: 9,011,219 URLs

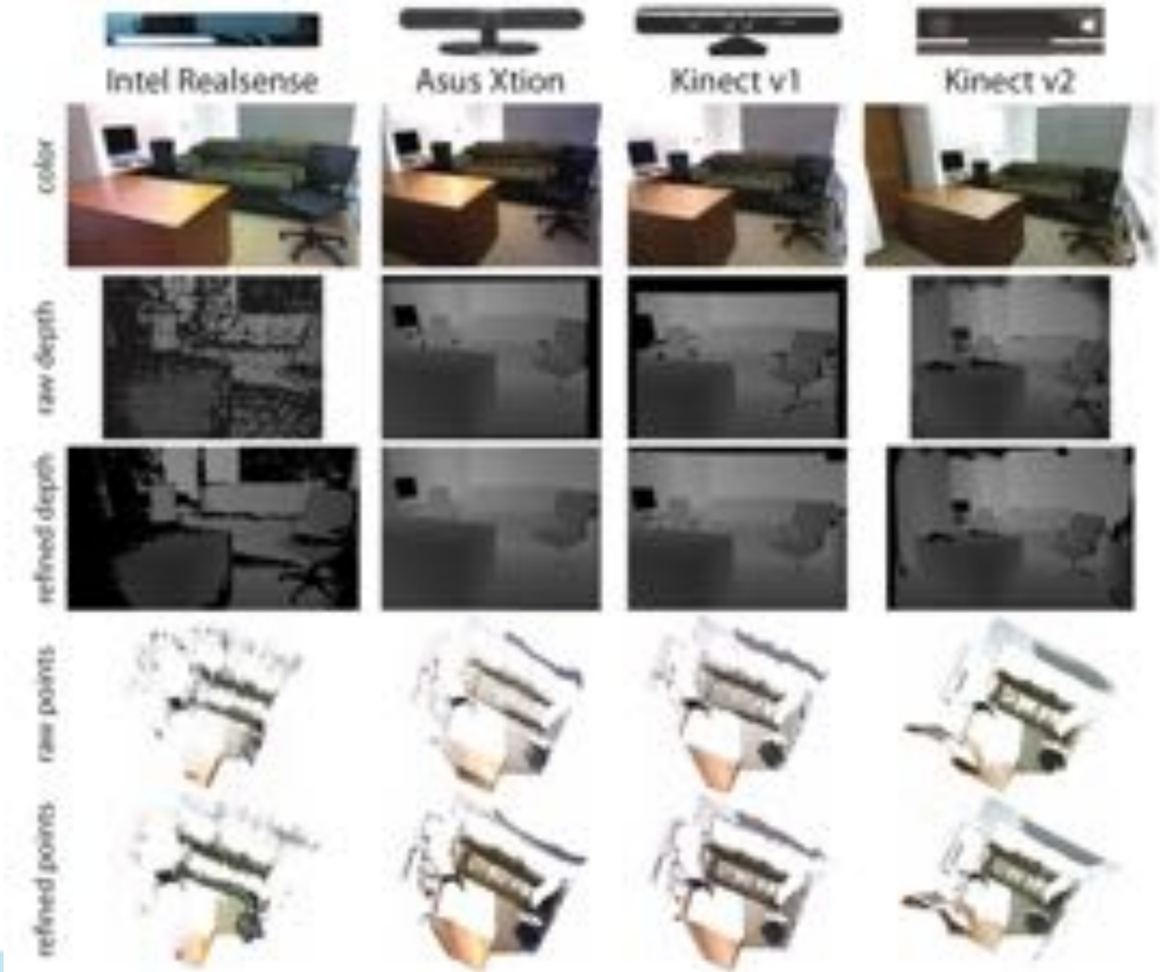
NUI Galway
OÉ Gaillimh

VIDEO DATASETS *(COMPRESSED!)*

YFCC100M	Large and diverse labeled image and video dataset	Flickr Videos and Images and associated description, titles, tags, and other metadata (such as EXIF and geotags)	100 million	**7 TB**
YouTube-8M	Large and diverse labeled video dataset	YouTube video IDs and associated labels from a diverse vocabulary of 4800 visual entities	8 million	**1.5 TB**

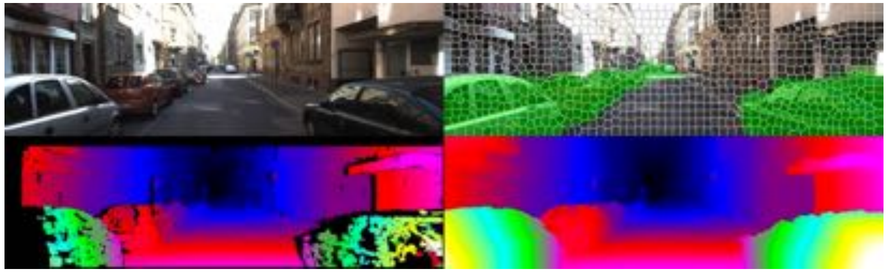
YOU NEED A GROUND TRUTH …

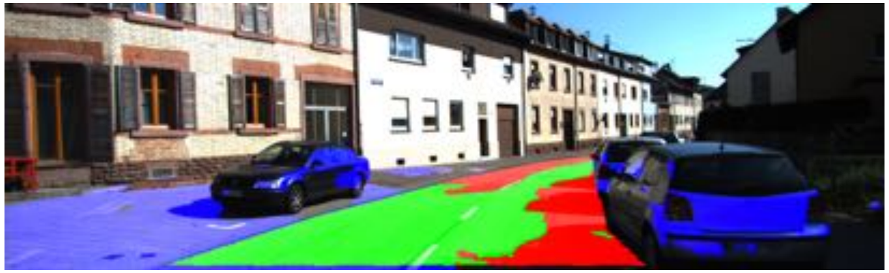
- Room views are an interesting data source for smart consumer devices …
- Determine the location of people and the acoustic properties of an ambient space …



NUI Galway
OÉ Gaillimh

FOR AUTONOMOUS VEHICLES …





Ground truth is
Challenging …

FOR HOME ENVIRONMENTS …

Figure 3. **Example images with annotation from our dataset.**

Sample results and cost considerations

The quality of the annotations provided by Mechanical Turk workers is in general quite good. The following are example annotations provided by the workers.

The following are statistics for the tasks that we submitted to Mechanical Turk.

Task	Price per image	Task description	# images	# annotations collected	Time elapsed (hours)	# workers
1	$0.01	Label at least one object in the image	237	678	13.5	37
2	$0.01	Label at least five objects in the image	271	1492	23.13	43
3	$0.01	Label as many objects as you wish in the image	271	627	9.08	28

YOU'LL NEED MANUAL ANNOTATIONS …

DRAWBACKS OF REAL DATA – *CHALLENGES INCLUDE: ACQUISITION, ANNOTATION & ERRORS*

- Acquisition [data + ground truth] & postprocessing costs
  - Challenge to get it all 'right', or you have to re-do everything
  - Data may depend on acquisition conditions (e.g. weather)
  - Data may depend on acquisition equipment (e.g. camera model_
- Annotation & Labeling Costs
  - Prone to human error; variable quality/skill of annotators
- Unquantifiable sources of data-error (noise, optical blur) that might impact on training features

THE 7TH **MAIN POINT** FOR TODAY!

- Real Data is Difficult ….

- … costly & complicated to Acquire, …
- … challenging to Annotate correctly …
- … prone to error & quality-control issues …
- **Get it wrong you have to do it all again!**

TOPIC #3 SOLVING THE PROBLEMS OF "REAL DATA" …

- Topic #1 – The Connection between Deep Learning & Computer Vision
  - Some Background on how I fell into the Deep Learning Pool
- Topic #2 – The Importance of Data for Deep Learning
  - What is Data Augmentation & why we need it
  - Smart-Augmentation – learning new ways to augment datasets
  - Boundary Equilibrius GAN (BEGAN) – a GAN to generate Face Images
  - Latent Spaces, learning Annotations & generating Custom Data
- **Topic #3 – Virtual Reality as a source of Datasets**
  - Real Data is difficult …
  - **Fake data is better … and the tools are here today!**

VIRTUAL REALITY IS NOW GOOD ENOUGH TO SIMULATE ‘REAL DATA’…

ULTRA-REALISTIC FACIAL ANIMATIONS

https://youtu.be/TxErDzsIdKI


Digital Version of Andy Serkis Recites 'Macbeth' to Show the Future of Performance Capture in Gaming

by Justin Page at 10:16 AM on March 23, 2018

VR ENVIRONMENTS FOR AUTONOMOUS VEHICLES & DRONES …

(AIRSIM FROM MICROSOFT RESEARCH …)

VR REAL WORLD ENVIRONMENTS WITH INCREDIBLE LEVELS OF DETAIL … RUNNING IN REAL-TIME ON A DESKTOP GPU NODE!
(“A BOY & HIS KITE DEMO ON UNREAL ENGINE 4…)

Epic’s Unreal Engine 4 - “A Boy and His Kite” featured physically-based rendering, with a single NVidia Titan-X GPU rendering ***100 square miles of terrain***.

It is complimented by ray-tranced distance field soft shadows, full scene HDR reflections, high quality motion blur and DOF, as well as distance field ambient occlusion.

The leaves and grass are all two-sided, with ***15 million pieces of vegetation***; photmetric sampling of real world data and dynamic global illumination from heightfields.

You really have to load this on a computer and explore it; there is ‘wind’ moving the leaves in the trees and rippling the surface of the water; you can load a plug-in for a drone or an SUV and explore this ‘World’ …

MANY BENEFITS OF 'FAKE DATA'…

- Provides a ***more accurate ground truth*** than real data, and of as many 3D points as needed …
- Large numbers of ***2D viewpoints*** can be rendered from a single 3D scene …
- Most ***annotations can be automated*** as part of the rendering process …
- Data is essentially ***free of noise & blur*** (but these can be simulated if needed) …
- For ***mixed reality*** we can incorporate acoustics & physics engines into the VR environment …

SOME LIMITATIONS …

- Generated Data is only as good as the 3D model/environment
  - e.g. modeling of human motion/activity may not be good enough
- Significant Computational Resources & Storage Requirements both for model and generated datasets …
- Time & Cost factors are still significant
  - But now invested in creating the VR Environment or 3D models rather than collecting and annotation data (tradeable assets?)

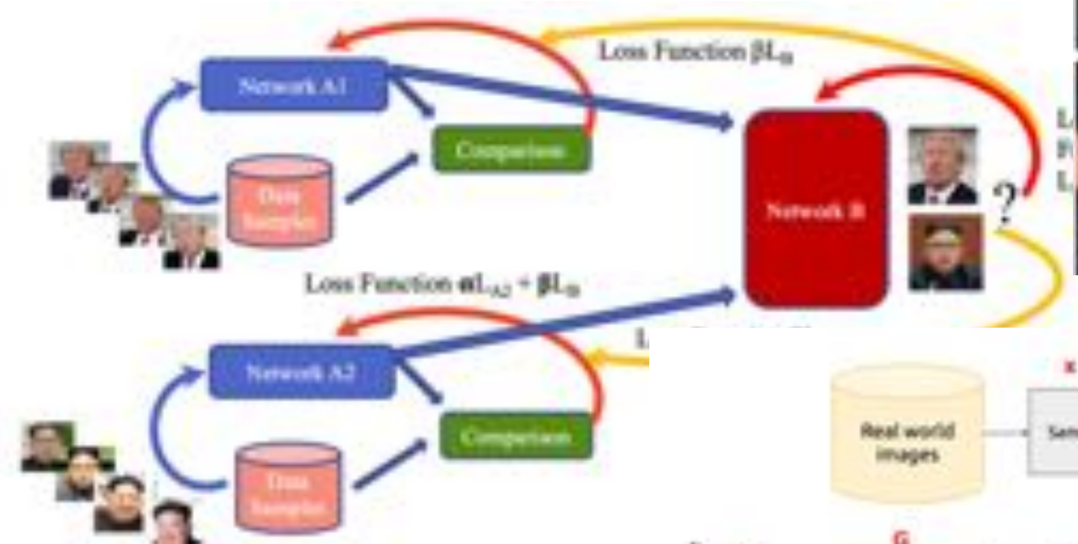
THE 8TH **MAIN POINT** FOR TODAY!

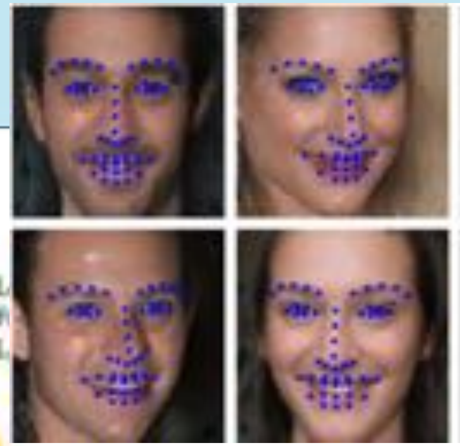
- **Fake Data is better than real ….**

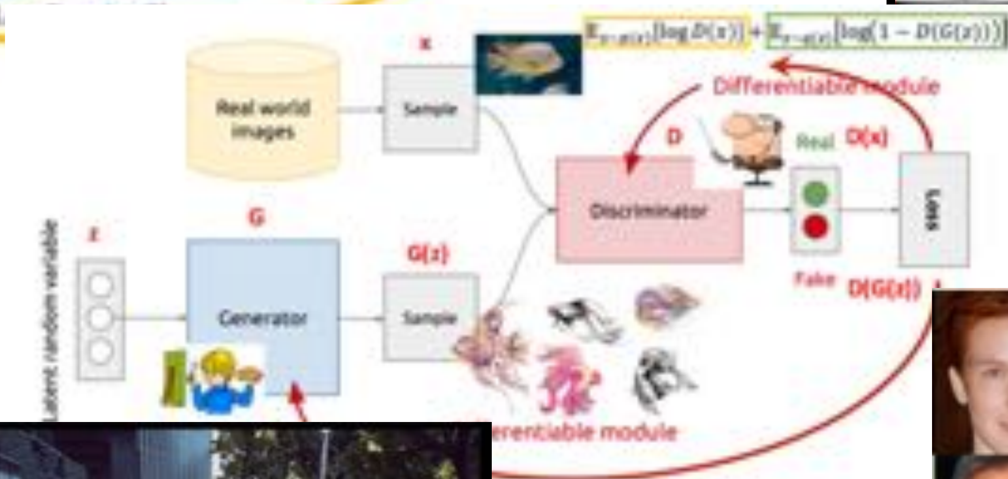
FINAL THOUGHTS

- We are at the 'tip' of an iceberg as researchers continue to refine these techniques and find new ways to scale, refine and develop these approaches …
- **"Fake Data"** has some interesting advantages:
    - Avoids ***privacy issues*** & side-steps ***new regulations*** (e.g. GDPR)
    - Significant cost reductions in *data acquisition* & *annotation* for very large datasets …
    - Solves challenges for *ground truth* and *multi-spectral registration* ...
    - Lots of potential to subtly control/refine how data is *generated*, *augmented* & *annotated* … we are still learning new 'tricks' here …
        - … an interesting new field for *research* & *experimentation* …
- Are we entering a new era of "***Generated Data***"? [post ***Big-Data*** era?]

THE AGE OF GENERATED DATA?







QUESTIONS?

RESOURCE SLIDE #1 – SUPPORTING PUBLICATIONS

- **General Deep Learning:**
    - Lemley J, Bazrafkan S, Corcoran P.; ***Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision.*** IEEE Consumer Electronics Magazine. 2017 Apr;6(2):48-56.
    - Bazrafkan S, Nedelcu T, Filipczuk P, Corcoran P.; ***Deep learning for facial expression recognition: A step closer to a smartphone that knows your moods.*** In Consumer Electronics (ICCE), 2017 IEEE International Conference on 2017 Jan 8 (pp. 217-220). IEEE.
    - Bazrafkan S, Corcoran P. ***Enhancing iris authentication on handheld devices using deep learning derived segmentation techniques.*** In Consumer Electronics (ICCE), 2018 IEEE International Conference on 2018 Jan 12 (pp. 1-2). IEEE.

RESOURCE SLIDE #2 – SUPPORTING PUBLICATIONS, CONTD.

- **Data Augmentation:**
  - J Lemley, S Bazrafkan, P Corcoran ***Smart Augmentation Learning an Optimal Data Augmentation Strategy*** IEEE ACCESS 5, 5858-5869. March 2017
  - J Lemley, S Bazrafkan, P Corcoran ***Transfer Learning of Temporal Information for Driver Action Classification*** Proceedings of the 28th Modern Artificial Intelligence and Cognitive Science Conference. April 2017
  - J Lemley, S Bazrafkan, P Corcoran ***Learning Data Augmentation for Consumer Devices and Services*** Consumer Electronics (ICCE), 2018 IEEE International Conference on. January 2018
  - Bazrafkan S, Thavalengal S, Corcoran P. ***An End to End Deep Neural Network for Iris Segmentation in Unconstrained Scenarios.*** arXiv preprint arXiv:1712.02877. 2017 Dec 7.

RESOURCE SLIDE #3 – SUPPORTING PUBLICATIONS, CONTD.

- **Generative Adversarial Networks:**
  - Bazrafkan S, Javidnia H, Corcoran P. ***Face Synthesis with Landmark Points from Generative Adversarial Networks and Inverse Latent Space Mapping.*** arXiv preprint arXiv:1802.00390. 2018 Feb 1.
  - Bazrafkan S, Javidnia H, Corcoran P. ***Versatile Auxiliary Classifier+ Generative Adversarial Network (VAC+ GAN); Training Conditional Generators.*** arXiv preprint arXiv:1805.00316. 2018 May 1.
- **Other Techniques (SPDNN, etc):**
  - Bazrafkan S, Javidnia H, Lemley J, Corcoran P. ***Depth from Monocular Images using a Semi-Parallel Deep Neural Network (SPDNN) Hybrid Architecture.*** arXiv preprint arXiv:1703.03867. 2017 Mar 10.
  - Bazrafkan S, Corcoran PM. ***Pushing the AI Envelope: Merging Deep Networks to Accelerate Edge Artificial Intelligence in Consumer Electronics Devices and Systems.*** IEEE Consumer Electronics Magazine. 2018 Mar;7(2):55-61.

RESOURCE SLIDE #4 – SUPPORTING PUBLICATIONS, CONTD.

- **Other Key Papers (Adversarial Networks, GANs, Boundary Equilibrium-GAN, etc):**
  - Nguyen A, Yosinski J, Clune J. ***Deep neural networks are easily fooled: High confidence predictions for unrecognizable images.*** In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015 (pp. 427-436).
  - Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. ***Generative adversarial nets.*** In Advances in neural information processing systems 2014 (pp. 2672-2680).
  - ***Generative Adversarial Networks (GANs)*** From Ian Goodfellow et al. A short tutorial by :- Binglin, Shashank & Bhargav
    - http://slazebni.cs.illinois.edu/spring17/lec11_gan.pdf
  - Berthelot D, Schumm T, Metz L. Began: Boundary equilibrium generative adversarial networks. arXiv preprint arXiv:1703.10717. 2017 Mar 31.

RESOURCE SLIDE #5 – SUPPORTING PUBLICATIONS, CONTD.

- **Other Key Papers (Adversarial Networks, GANs, Boundary Equilibrium-GAN, etc):**

  - Berthelot D, Schumm T, Metz L. Began: Boundary equilibrium generative adversarial networks. arXiv preprint arXiv:1703.10717. 2017 Mar 31.
  - Rosca M, Lakshminarayanan B, Warde-Farley D, Mohamed S. Variational approaches for auto-encoding generative adversarial networks. arXiv preprint arXiv:1706.04987. 2017 Jun 15.
  - Huang B, Chen W, Wu X, Lin CL, Suganthan PN. High-Quality Face Image Generated with Conditional Boundary Equilibrium Generative Adversarial Networks. Pattern Recognition Letters. 2018 Apr 19.
  - Asthana A, Zafeiriou S, Cheng S, Pantic M. Incremental face alignment in the wild. InProceedings of the IEEE conference on computer vision and pattern recognition 2014 (pp. 1859-1866).
  - Chrysos GG, Antonakos E, Snape P, Asthana A, Zafeiriou S. A comprehensive performance evaluation of deformable face tracking “in-the-wild”. International Journal of Computer Vision. 2018 Apr 1;126(2-4):198-232.
  - Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. arXiv:1412.6572. 2014 Dec 20.
  - Su J, Vargas DV, Kouichi S. One pixel attack for fooling deep neural networks. arXiv preprint arXiv:1710.08864. 2017 Oct 24.