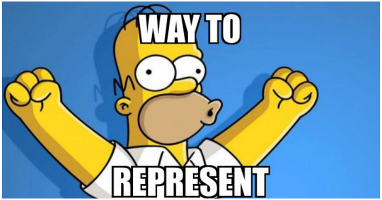
REALLY!

REALLY!

REALLY!



Source: F.R.I.E.N.D.S , www.makeameme.com

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Advances in Representation Learning

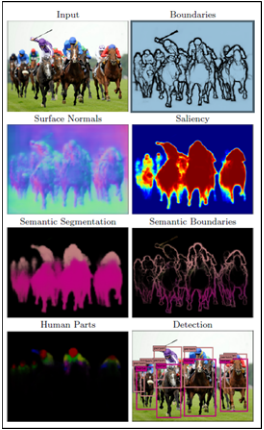
"representing NN in a fancy way!"

Arijit Ghosh
Seminar Advanced Deep Learning, Friedrich-Alexander-Universität Erlangen-Nürnberg
January 25, 2023

DeepNets are "Amazing" 101

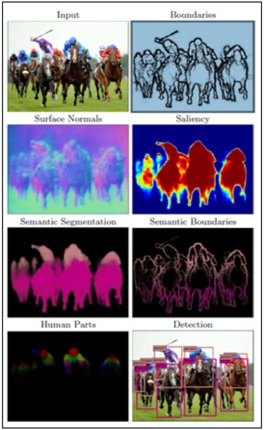
Kokkinos, Iasonas. "Ubernet: Training a universal convolutional neural network for low-, mid-, and high-level vision using diverse datasets and limited memory." CVPR, 2017.

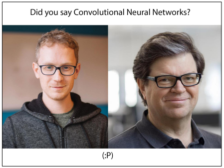
DeepNets are "Amazing" 101



Kokkinos, Iasonas. "Ubernet: Training a universal convolutional neural network for low-, mid-, and high-level vision using diverse datasets and limited memory." CVPR, 2017.

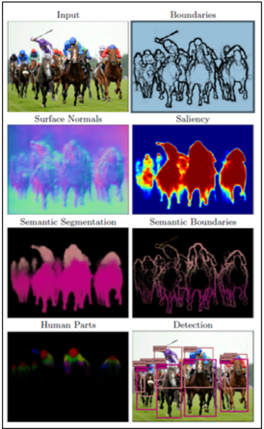
DeepNets are "Amazing" 101



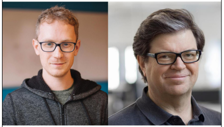


Kokkinos, Iasonas. "Ubernet: Training a universal convolutional neural network for low-, mid-, and high-level vision using diverse datasets and limited memory." CVPR, 2017.

DeepNets are "Amazing" 101



Did you say Convolutional Neural Networks?



(:P)

Kokkinos, Iasonas. "Ubernet: Training a universal convolutional neural network for low-, mid-, and high-level vision using diverse datasets and limited memory." CVPR, 2017.

DeepNets are "Amazing" 102.... Really ?

Carter, Brandon, et al. "Overinterpretation reveals image classification model pathologies." NeurIPS, 2021.

DeepNets are "Amazing" 102.... Really ?

Carter, Brandon, et al. "Overinterpretation reveals image classification model pathologies." NeurIPS, 2021.

DeepNets are "Amazing" 102.... Really ?

ResNet, AlexNet, VGG Accuracy $\geq$ 92%

Carter, Brandon, et al. "Overinterpretation reveals image classification model pathologies." NeurIPS, 2021.

DeepNets are "Amazing" 102.... Really ?

NO

Yuille, Alan L., and Chenxi Liu. "Deep nets: What have they ever done for vision?." International Journal of Computer Vision 129.3 (2021)

DeepNets are "Amazing" 102.... Really ?

NO

➔ Problem of **Generalization**.

Yuille, Alan L., and Chenxi Liu. "Deep nets: What have they ever done for vision?." International Journal of Computer Vision 129.3 (2021)

DeepNets are "Amazing" 102.... Really ?

NO

➔ Problem of **Generalization**.

ImageNet Image Recognition

Yuille, Alan L., and Chenxi Liu. "Deep nets: What have they ever done for vision?." International Journal of Computer Vision 129.3 (2021)

DeepNets are "Amazing" 102.... Really ?

NO

➔ Problem of **Generalization**.

ImageNet Image Recognition

Detection of Human Joints

in Leeds Sports Dataset

Yuille, Alan L., and Chenxi Liu. "Deep nets: What have they ever done for vision?." International Journal of Computer Vision 129.3 (2021)

DeepNets are "Amazing" 102.... Really ?

NO

➔ Problem of **Generalization**.

ImageNet Image Recognition

Detection of Human Joints in Leeds Sports Dataset

Yuille, Alan L., and Chenxi Liu. "Deep nets: What have they ever done for vision?." International Journal of Computer Vision 129.3 (2021)

DeepNets are "Amazing" 102.... Really ?

NO

➔ Problem of **Generalization**.

ImageNet Image Recognition

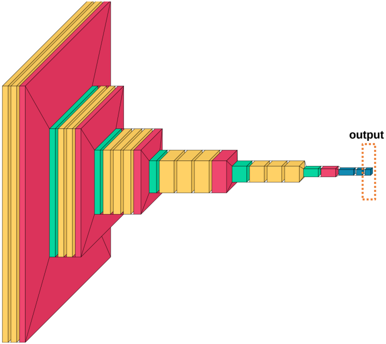
Detection of Human Joints in Leeds Sports Dataset

Yuille, Alan L., and Chenxi Liu. "Deep nets: What have they ever done for vision?." International Journal of Computer Vision 129.3 (2021)

Representation Learning vs Deep Learning - FIGHT!

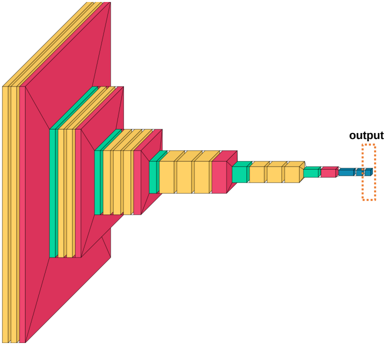
Representation Learning vs Deep Learning - FIGHT!

Common DL Models



Representation Learning vs Deep Learning - FIGHT!

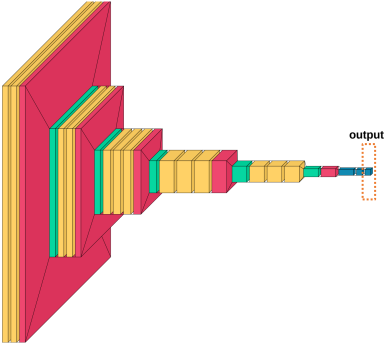
Common DL Models



➔ Specific **Targets**.

Representation Learning vs Deep Learning - FIGHT!

Common DL Models

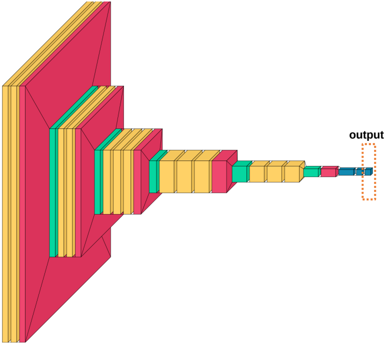


➔ Specific **Targets**.

Representation Learning

Representation Learning vs Deep Learning - FIGHT!

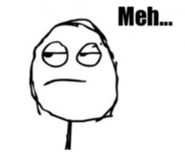
Common DL Models



➔ Specific **Targets**.

Representation Learning

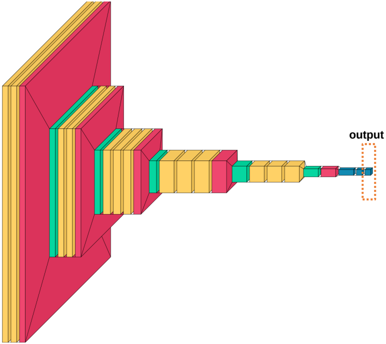
Specific Outputs?



Source: https://forum.game-labs.net/gallery/image/2682-meh-meme/

Representation Learning vs Deep Learning - FIGHT!

Common DL Models



➔ Specific **Targets**.

Representation Learning

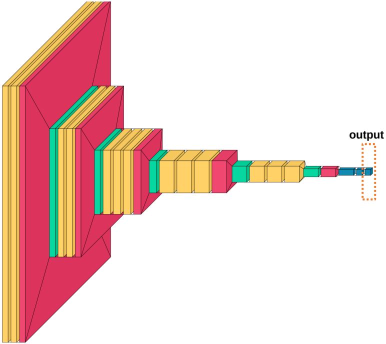
MY PRECIOUS...

FEATURES

Source: Lord of the Rings

Representation Learning ~~vs~~ Deep Learning - ~~FIGHT~~!

Common DL Models



➔ Specific **Targets**.

Representation Learning

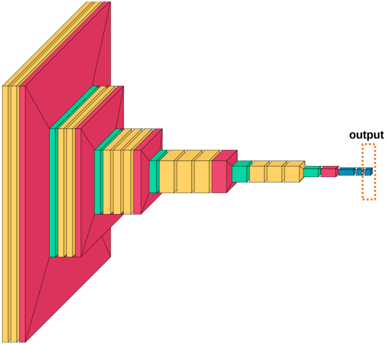
MY PRECIOUS…

FEATURES

Source: Lord of the Rings

Representation Learning with Deep Learning - CONQUER!

Common DL Models



➔ Specific **Targets**.

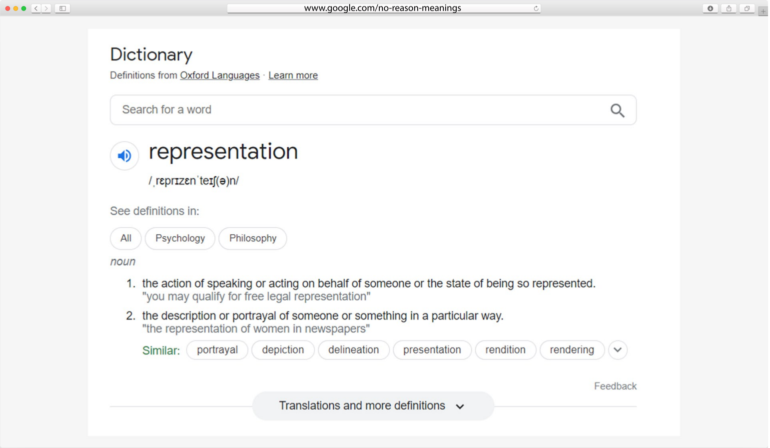
Representation Learning

MY PRECIOUS...

FEATURES

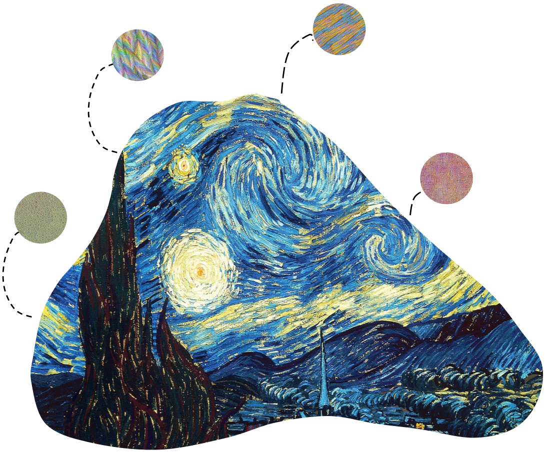
Source: Lord of the Rings

Motivating Representation Learning 101



Source: www.google.com

Motivating Representation Learning 101 (ctd...)



Source: (Van Gogh, The Starry Night)

Motivating Representation Learning 102

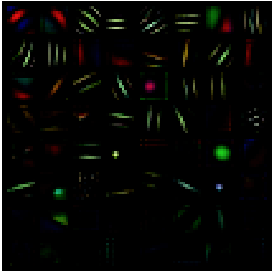
Visualization of First Layer weights

Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised representation learning by predicting image rotations." arXiv preprint, 2018.

https://cs231n.github.io/convolutional-networks/

Motivating Representation Learning 102

Visualization of First Layer weights



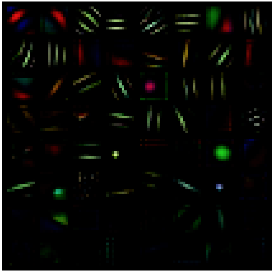
AlexNet on CIFAR-10 Object Recognition Task

Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised representation learning by predicting image rotations." arXiv preprint, 2018.

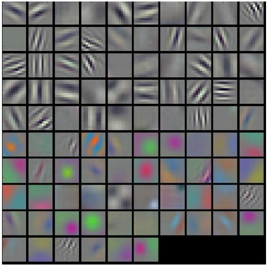
`https://cs231n.github.io/convolutional-networks/`

Motivating Representation Learning 102

Visualization of First Layer weights



AlexNet on CIFAR-10 Object Recognition Task



AlexNet on Imagenet Object Recognition Task

Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised representation learning by predicting image rotations." arXiv preprint, 2018.

https://cs231n.github.io/convolutional-networks/

Outline

Source: F.R.I.E.N.D.S

Hi, Clustering !

Source: https://www.geeksforgeeks.org/clustering-in-machine-learning/

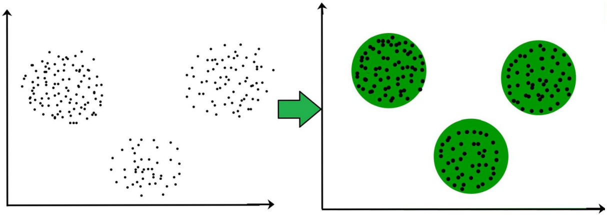
Hi, Clustering !

➔ Learning **Data Manifolds**.

Source: `https://www.geeksforgeeks.org/clustering-in-machine-learning/`

Hi, Clustering !

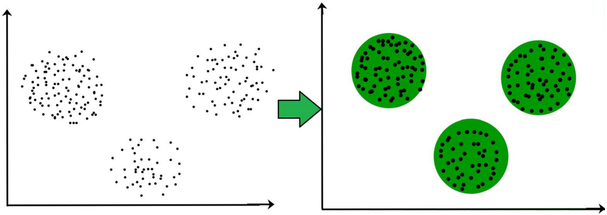
➔ Learning **Data Manifolds**.



Source: https://www.geeksforgeeks.org/clustering-in-machine-learning/

Hi, Clustering !

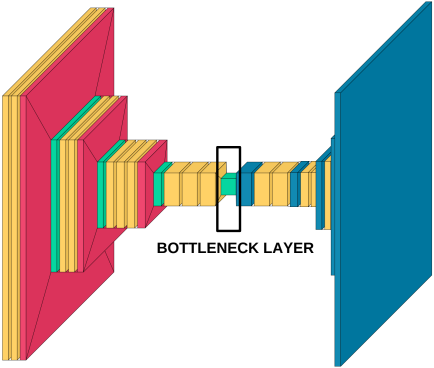
➔ Learning **Data Manifolds**.



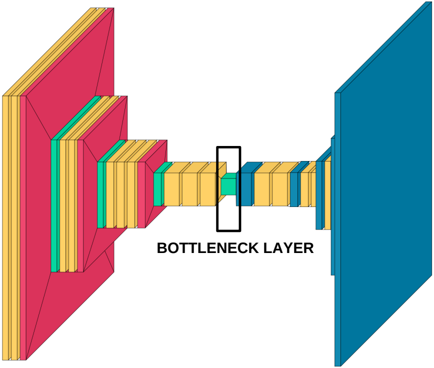
➔ Clustering with Neural Networks?

Source: `https://www.geeksforgeeks.org/clustering-in-machine-learning/`

Deep Clustering 101

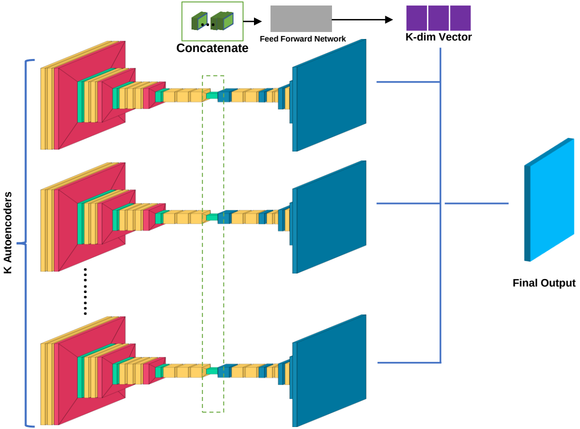


Deep Clustering 101



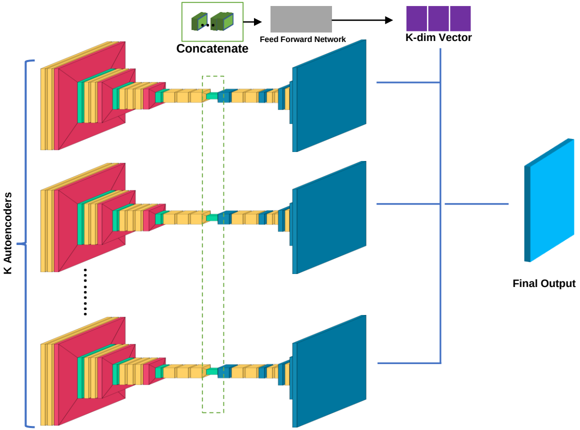
➔ Highly tangled.

Deep Clustering 102



Zhang, Dejiao, et al. "Deep unsupervised clustering using mixture of autoencoders." arXiv preprint, 2017.

Deep Clustering 102



➔ Disentangled.

Zhang, Dejiao, et al. "Deep unsupervised clustering using mixture of autoencoders." arXiv preprint, 2017.

Deep Clustering 102 : The Problem!

Deep Clustering 102 : I hate "K"

Deep Clustering 103 : Intoducing DeepDPM

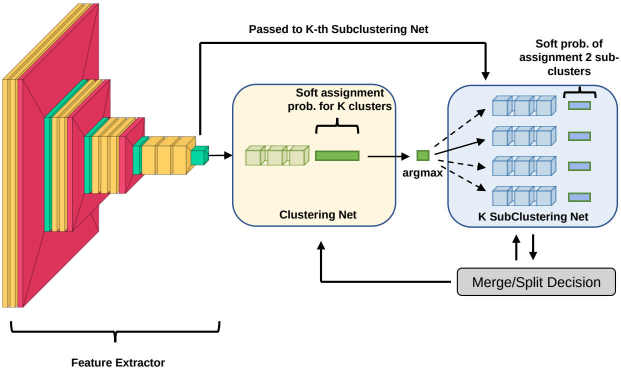
Ronen, Meitar, Shahaf E. Finder, and Oren Freifeld. "DeepDPM: Deep Clustering With an Unknown Number of Clusters." CVPR, 2022.

Deep Clustering 103 : Intoducing DeepDPM

➔ No dependency on K.

Ronen, Meitar, Shahaf E. Finder, and Oren Freifeld. "DeepDPM: Deep Clustering With an Unknown Number of Clusters." CVPR, 2022.

Deep Clustering 103 : Intoducing DeepDPM



Ronen, Meitar, Shahaf E. Finder, and Oren Freifeld. "DeepDPM: Deep Clustering With an Unknown Number of Clusters." CVPR, 2022.

Deep Clustering: A reality check

Ronen, Meitar, Shahaf E. Finder, and Oren Freifeld. "DeepDPM: Deep Clustering With an Unknown Number of Clusters." CVPR, 2022.

Deep Clustering: A reality check

- State-of-the-art **clustering results**.

Ronen, Meitar, Shahaf E. Finder, and Oren Freifeld. "DeepDPM: Deep Clustering With an Unknown Number of Clusters." CVPR, 2022.

Deep Clustering: A reality check

- State-of-the-art **clustering results**.
- **Comparable** to **Supervised Nets**?

Ronen, Meitar, Shahaf E. Finder, and Oren Freifeld. "DeepDPM: Deep Clustering With an Unknown Number of Clusters." CVPR, 2022.

Deep Clustering: A reality check

- State-of-the-art **clustering results**.
- **Comparable** to **Supervised Nets**?

NO

Ronen, Meitar, Shahaf E. Finder, and Oren Freifeld. "DeepDPM: Deep Clustering With an Unknown Number of Clusters." CVPR, 2022.

Introducing Self-Supervised Learning

Introducing Self-Supervised Learning

- Bring Your Own Labels.

Introducing Self-Supervised Learning

- Bring Your Own Labels.
- Proxy-Tasks.

Introducing Self-Supervised Learning

- Bring Your Own Labels.
- Proxy-Tasks.

WATCHING BREAKING BAD

BEFORE CHEMISTRY EXAM

Source: Breaking Bad

Proxy Task 101 : Rotation with RotNet

Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised representation learning by predicting image rotations." arXiv preprint, 2018.

Proxy Task 101 : Rotation with RotNet

- **Labels** based on **Image Rotation**

Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised representation learning by predicting image rotations." arXiv preprint, 2018.

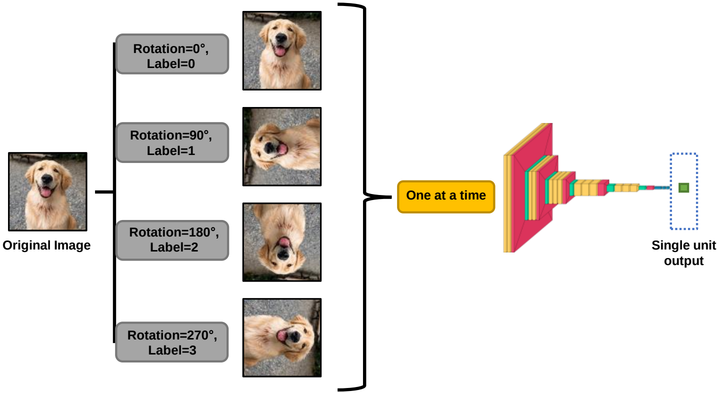
Proxy Task 101 : Rotation with RotNet

- **Labels** based on **Image Rotation** .
- 4 labels.

Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised representation learning by predicting image rotations." arXiv preprint, 2018.

Proxy Task 101 : Rotation with RotNet

- **Labels** based on **Image Rotation** .
- 4 labels.

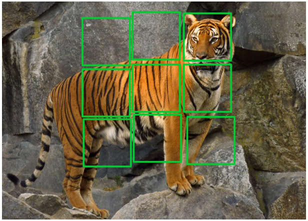


Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised representation learning by predicting image rotations." arXiv preprint, 2018.

Proxy Task 102 : Puzzle Solving

Noroozi, Mehdi, and Paolo Favaro. "Unsupervised learning of visual representations by solving jigsaw puzzles." ECCV, 2016.

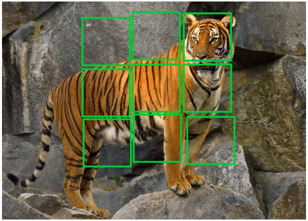
Proxy Task 102 : Puzzle Solving



Input Image

Noroozi, Mehdi, and Paolo Favaro. "Unsupervised learning of visual representations by solving jigsaw puzzles." ECCV, 2016.

Proxy Task 102 : Puzzle Solving

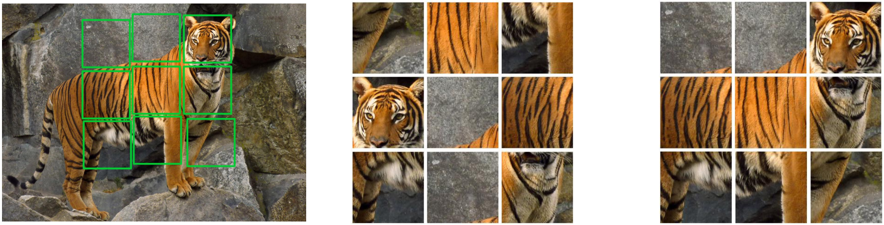


Input Image

Random Cropped and Shuffled as a Puzzle

Noroozi, Mehdi, and Paolo Favaro. "Unsupervised learning of visual representations by solving jigsaw puzzles." ECCV, 2016.

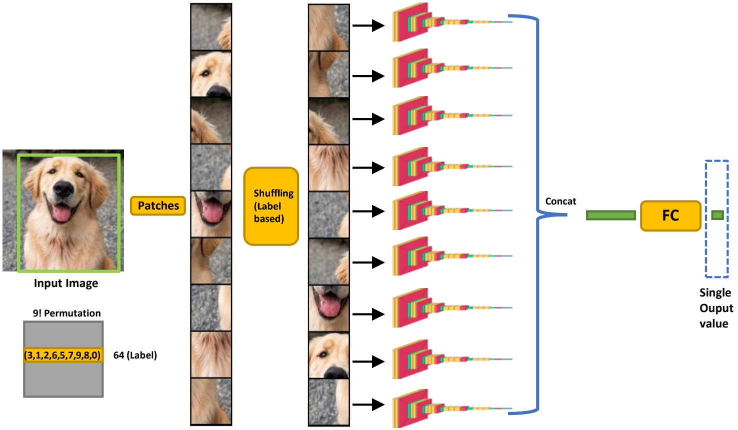
Proxy Task 102 : Puzzle Solving



Input Image | Random Cropped and Shuffled as a Puzzle | Solved Puzzle

Noroozi, Mehdi, and Paolo Favaro. "Unsupervised learning of visual representations by solving jigsaw puzzles." ECCV, 2016.

Proxy Task 102 : Puzzle Solving with Context Free Network



Noroozi, Mehdi, and Paolo Favaro. "Unsupervised learning of visual representations by solving jigsaw puzzles." ECCV, 2016.

Proxy Task 103 : Colorization

Zhang, Richard, Phillip Isola, and Alexei A. Efros. "Colorful image colorization." , ECCV, 2016.

Proxy Task 103 : Colorization

➔ Grayscale to Color Conversion.

Zhang, Richard, Phillip Isola, and Alexei A. Efros. "Colorful image colorization." , ECCV, 2016.

Proxy Task 103 : Colorization

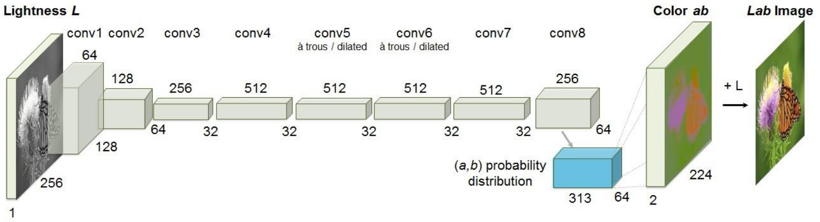
➔ Grayscale to Color Conversion.

Input Image

Output Image

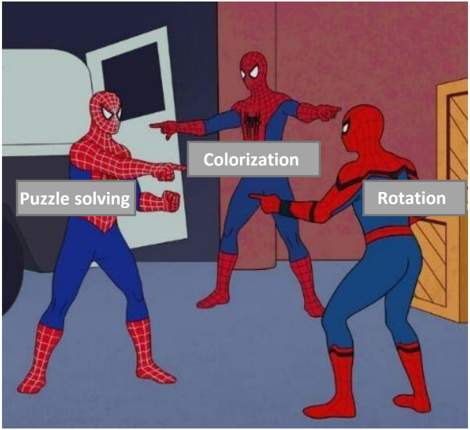
Zhang, Richard, Phillip Isola, and Alexei A. Efros. "Colorful image colorization." , ECCV, 2016.

Proxy Task 103 : Colorization with FCN



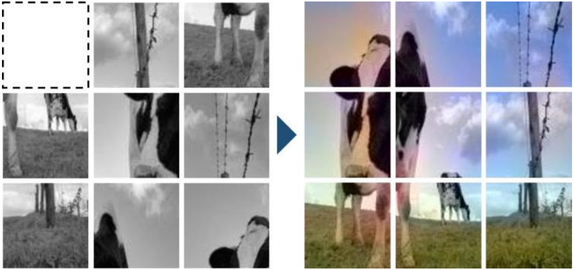
Zhang, Richard, Phillip Isola, and Alexei A. Efros. "Colorful image colorization." , ECCV, 2016.

Can we combine them?



Source: https://www.bradleyscout.com/voice/rumored-three-spider-mans-in-one/

Proxy Task 104: Mixture of all

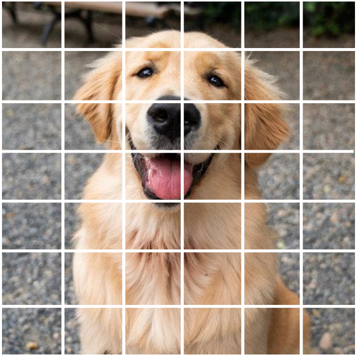


Kim, Dahun, Donghyeon Cho, Donggeun Yoo, and In So Kweon. "Learning image representations by completing damaged jigsaw puzzles." WACV, 2018.

Proxy Task 105: Efficient Masking

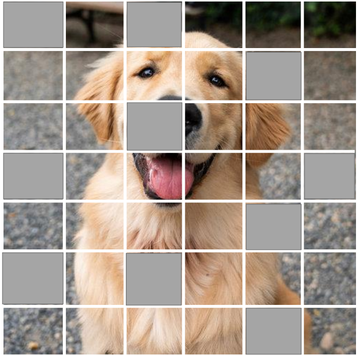
He, Kaiming, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. "Masked autoencoders are scalable vision learners." CVPR, 2022.

Proxy Task 105: Efficient Masking



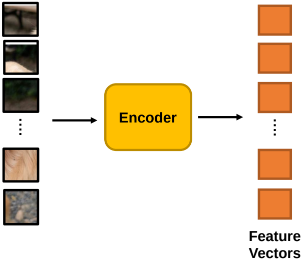
He, Kaiming, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. "Masked autoencoders are scalable vision learners." CVPR, 2022.

Proxy Task 105: Efficient Masking



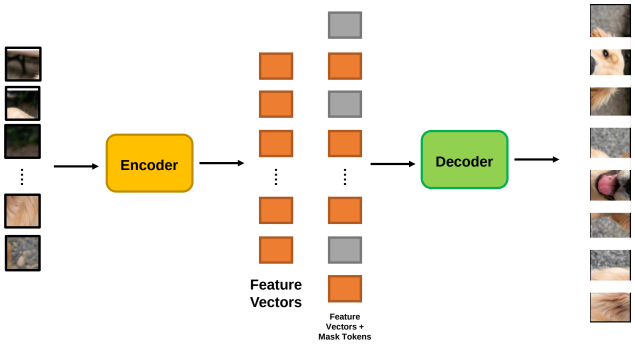
He, Kaiming, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. "Masked autoencoders are scalable vision learners." CVPR, 2022.

Proxy Task 105: Efficient Masking



He, Kaiming, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. "Masked autoencoders are scalable vision learners." CVPR, 2022.

Proxy Task 105: Efficient Masking



He, Kaiming, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. "Masked autoencoders are scalable vision learners." CVPR, 2022.

A bit is better than none :D

Bonjour, Semi-Supervised Learning

Bonjour, Semi-Supervised Learning

- Small portion **Labeled**.

Bonjour, Semi-Supervised Learning

- Small portion **Labeled**.
- How to take advantage?

Bonjour, Semi-Supervised Learning

- Small portion **Labeled**.
- How to take advantage?

I AM LEARNING

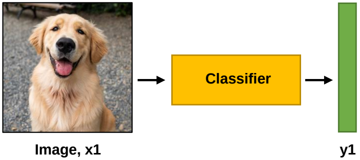
REPRESENTATIONS

Source: `https://openai.com/dall-e-2/`

Consistency Regularization

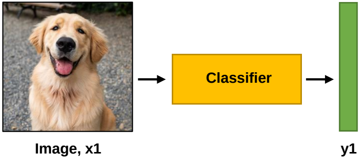
Tarvainen, Antti, and Harri Valpola. "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results." NeurIPS , 2017.

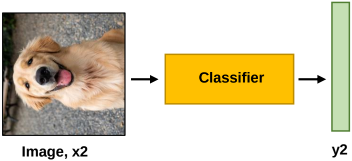
Consistency Regularization



Tarvainen, Antti, and Harri Valpola. "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results." NeurIPS , 2017.

Consistency Regularization





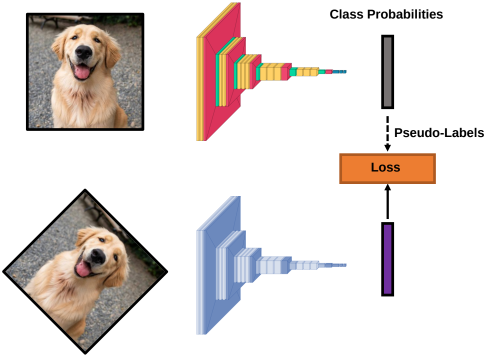
Tarvainen, Antti, and Harri Valpola. "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results." NeurIPS , 2017.

Consistency Regularization

$$y1 \approx y2$$

Tarvainen, Antti, and Harri Valpola. "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results." NeurIPS , 2017.

Pseudo Labels



Lee, Dong-Hyun. "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks." Workshop on challenges in representation learning, ICML , 2013.

MixUp

Zhang, Hongyi, et al. "mixup: Beyond empirical risk minimization." arXiv preprint , 2017.

MixUp

image_1

Label = [0 , 1]

Zhang, Hongyi, et al. "mixup: Beyond empirical risk minimization." arXiv preprint , 2017.

MixUp

image_1

Label = [0 , 1]

image_2

Label = [1 , 0]

Zhang, Hongyi, et al. "mixup: Beyond empirical risk minimization." arXiv preprint , 2017.

MixUp

image_1

Label = [0 , 1]

image_2

Label = [1 , 0]

Mixed_Image = λ * image_1 + (1 - λ) * image_2

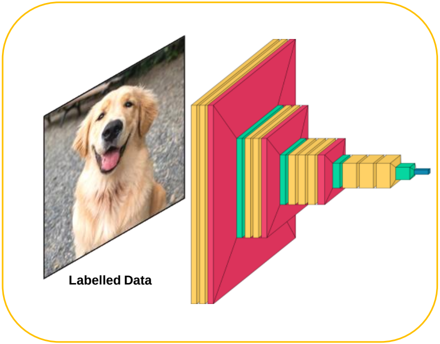
λ = 0.4 , Label = [0.6 , 0.4]

Zhang, Hongyi, et al. "mixup: Beyond empirical risk minimization." arXiv preprint , 2017.

MixMatch : Combining them ALL

Berthelot, David, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A. Raffel. "Mixmatch: A holistic approach to semi-supervised learning." NeurIPS , 2019).

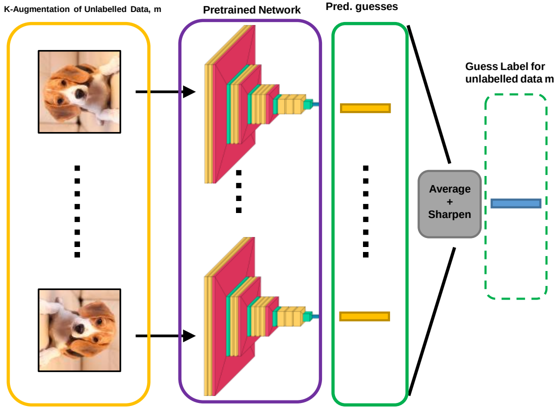
MixMatch : Combining them ALL



Pretraining

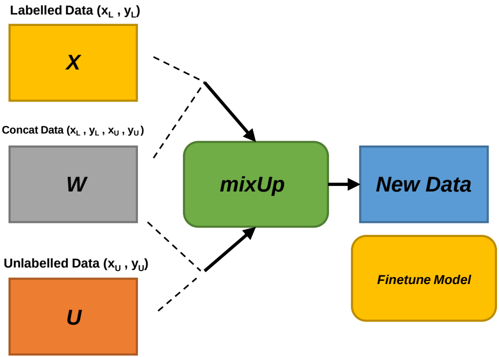
Berthelot, David, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A. Raffel. "Mixmatch: A holistic approach to semi-supervised learning." NeurIPS , 2019).

MixMatch : Combining them ALL



Berthelot, David, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A. Raffel. "Mixmatch: A holistic approach to semi-supervised learning." NeurIPS , 2019).

MixMatch : Combining them ALL



Berthelot, David, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A. Raffel. "Mixmatch: A holistic approach to semi-supervised learning." NeurIPS , 2019).

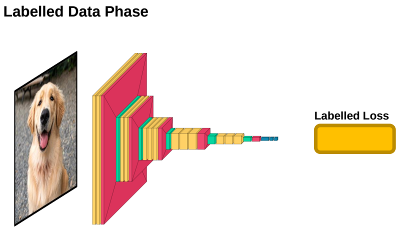
The thing is...

Source: `https://makeameme.org/meme/you-were-supposed-882fc96e1a`

Fixed by FixMatch

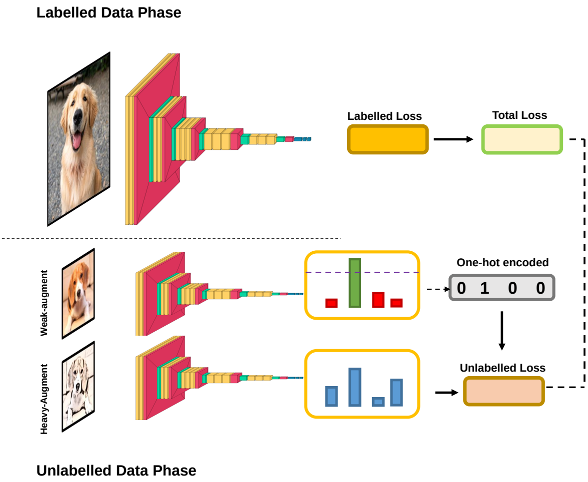
Sohn, Kihyuk, David Berthelot, Nicholas Carlini, Zizhao Zhang, Han Zhang, Colin A. Raffel, Ekin Dogus Cubuk, Alexey Kurakin, and Chun-Liang Li. "Fixmatch: Simplifying semi-supervised learning with consistency and confidence." NeurIPS , 2020.

Fixed by FixMatch



Sohn, Kihyuk, David Berthelot, Nicholas Carlini, Zizhao Zhang, Han Zhang, Colin A. Raffel, Ekin Dogus Cubuk, Alexey Kurakin, and Chun-Liang Li. "Fixmatch: Simplifying semi-supervised learning with consistency and confidence." NeurIPS , 2020.

Fixed by FixMatch



Sohn, Kihyuk, David Berthelot, Nicholas Carlini, Zizhao Zhang, Han Zhang, Colin A. Raffel, Ekin Dogus Cubuk, Alexey Kurakin, and Chun-Liang Li. "Fixmatch: Simplifying semi-supervised learning with consistency and confidence." NeurIPS , 2020.

But....

But....

➔ Explicit need of Features in Supervised Learning.

But....

➔ Explicit need of Features in Supervised Learning.

METRIC LEARNING

Contrastive Loss

Hadsell, Raia, Sumit Chopra, and Yann LeCun. "Dimensionality reduction by learning an invariant mapping." CVPR 2006

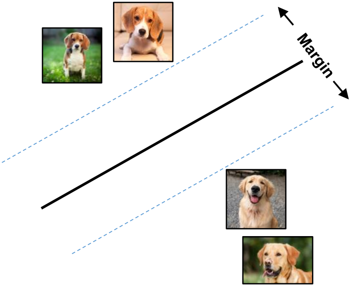
Contrastive Loss

➔ Learn an embedding/ feature vector.

Hadsell, Raia, Sumit Chopra, and Yann LeCun. "Dimensionality reduction by learning an invariant mapping." CVPR 2006

Contrastive Loss

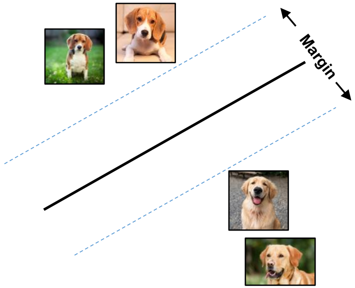
➔ Learn an embedding/ feature vector.



Hadsell, Raia, Sumit Chopra, and Yann LeCun. "Dimensionality reduction by learning an invariant mapping." CVPR 2006

Contrastive Loss

➔ Learn an embedding/ feature vector.



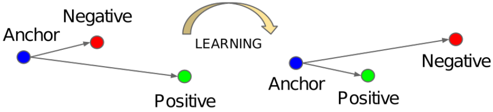
$$L_{contrastive} = \mathbb{1}_{y_1=y_2} D^2 f_\theta(x_1, x_2) + \mathbb{1}_{y_1 \neq y_2} max(0, \alpha - D^2 f_\theta(x_1, x_2))$$

Hadsell, Raia, Sumit Chopra, and Yann LeCun. "Dimensionality reduction by learning an invariant mapping." CVPR 2006

Triplet Loss

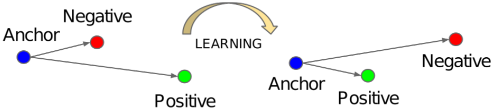
Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering." CVPR, 2015.

Triplet Loss



Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering." CVPR, 2015.

Triplet Loss



$$L_{triplet} = max(0, \alpha + D^2 f_\theta(x_a, x_p) - D^2 f_\theta(x_a, x_n))$$

Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering." CVPR, 2015.

Triplet Loss... The Problems

Source: https://hav4ik.github.io/articles/deep-metric-learning-survey

Triplet Loss... The Problems

- **Expansion** Problem.

Source: https://hav4ik.github.io/articles/deep-metric-learning-survey

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Triplet Loss... The Problems

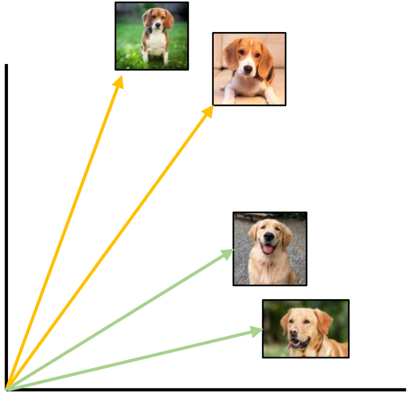
- **Expansion** Problem.
- **Sampling** Problem.

Source: https://hav4ik.github.io/articles/deep-metric-learning-survey

Introducing... SphereFace

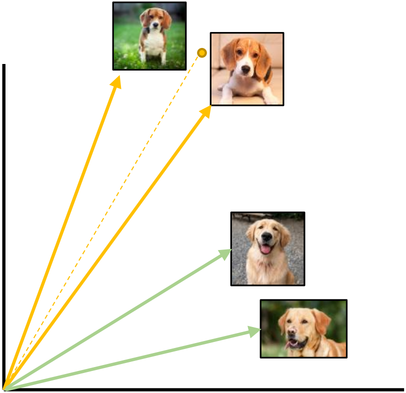
Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

Introducing... SphereFace



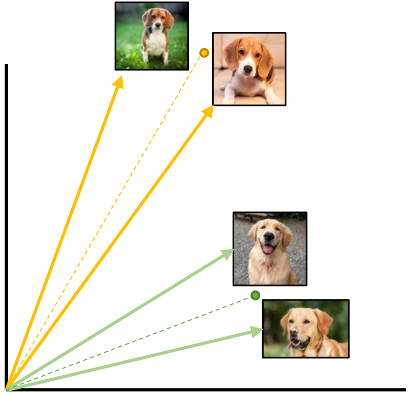
Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

Introducing... SphereFace



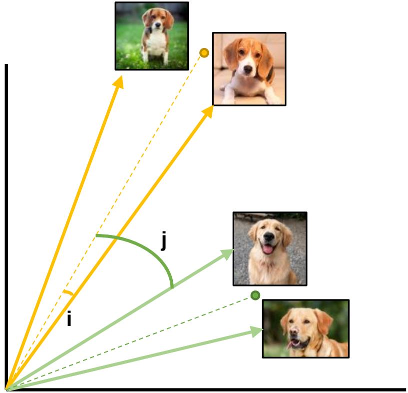
Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

Introducing... SphereFace



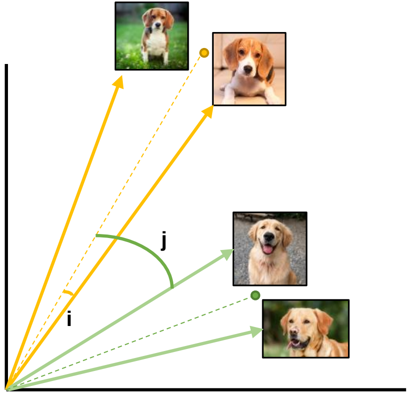
Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

Introducing... SphereFace



Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

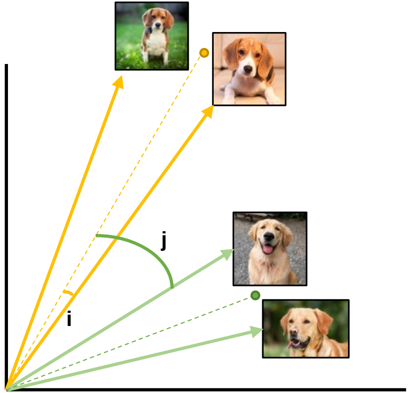
Introducing... SphereFace



- i < j ➔ Class i.

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

Introducing... SphereFace



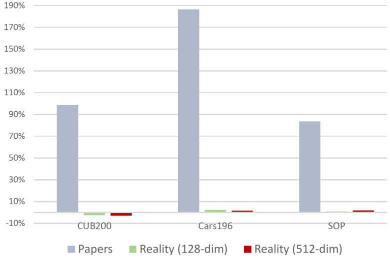
- i < j ➔ Class i.
- Makes **last layer weights** as **class centers**.

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

Don’t believe in what you see...

Musgrave, Kevin, Serge Belongie, and Ser-Nam Lim. "A metric learning reality check." ECCV, 2020.

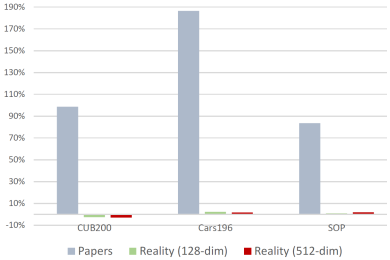
Don’t believe in what you see...



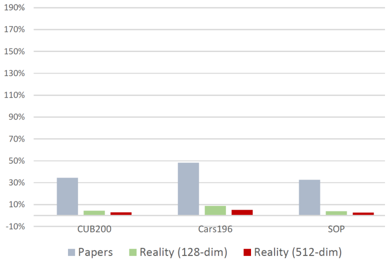
Improvements over Contrastive Loss

Musgrave, Kevin, Serge Belongie, and Ser-Nam Lim. "A metric learning reality check." ECCV, 2020.

Don’t believe in what you see...



Improvements over Contrastive Loss

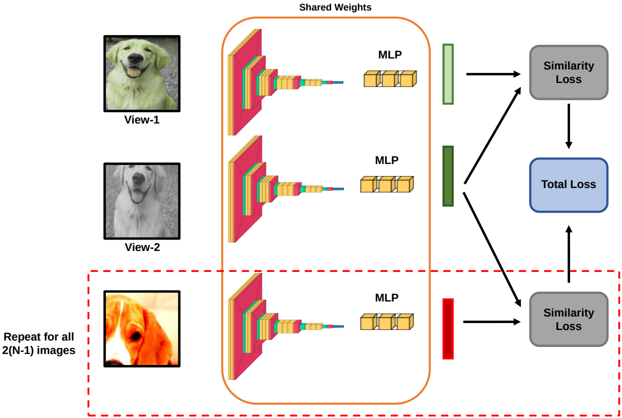


Improvements over Triplet Loss

Musgrave, Kevin, Serge Belongie, and Ser-Nam Lim. "A metric learning reality check." ECCV, 2020.

Hybrid Methods 101... SimCLR

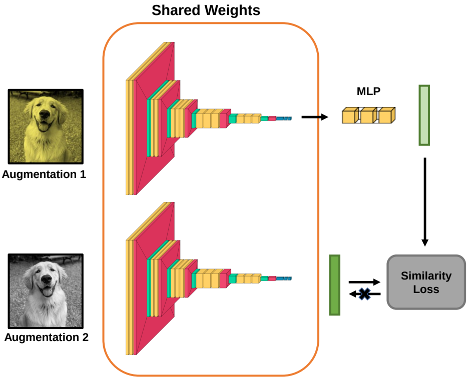
➔ Self-Supervised + Contrastive Learning



Chen, Ting, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. "A simple framework for contrastive learning of visual representations." PMLR, 2020.

Hybrid Methods 102... SimSiam

➔ Self-Supervised + Contrastive Learning



Chen, Xinlei, and Kaiming He. "Exploring simple siamese representation learning." CVPR, 2021.

Conclusion...

Koppula, Skanda, Yazhe Li, Evan Shelhamer, Andrew Jaegle, Nikhil Parthasarathy, Relja Arandjelovic, João Carreira, and Olivier Hénaff. "Where should i spend my flops? efficiency evaluations of visual pre-training methods." arXiv preprint arXiv:2209.15589 (2022).

Conclusion...

- Representation Learning methods- Amazing!

Koppula, Skanda, Yazhe Li, Evan Shelhamer, Andrew Jaegle, Nikhil Parthasarathy, Relja Arandjelovic, João Carreira, and Olivier Hénaff. "Where should i spend my flops? efficiency evaluations of visual pre-training methods." arXiv preprint arXiv:2209.15589 (2022).

Conclusion...

- Representation Learning methods- Amazing!
- Pre-Training.

Koppula, Skanda, Yazhe Li, Evan Shelhamer, Andrew Jaegle, Nikhil Parthasarathy, Relja Arandjelovic, João Carreira, and Olivier Hénaff. "Where should i spend my flops? efficiency evaluations of visual pre-training methods." arXiv preprint arXiv:2209.15589 (2022).

Conclusion...

- Representation Learning methods- Amazing!
- Pre-Training.
- CO_2 foot print.

Koppula, Skanda, Yazhe Li, Evan Shelhamer, Andrew Jaegle, Nikhil Parthasarathy, Relja Arandjelovic, João Carreira, and Olivier Hénaff. "Where should i spend my flops? efficiency evaluations of visual pre-training methods." arXiv preprint arXiv:2209.15589 (2022).

Conclusion...

- Representation Learning methods- Amazing!
- Pre-Training.
- CO_2 foot print.
- Better Dataset Curation.

Koppula, Skanda, Yazhe Li, Evan Shelhamer, Andrew Jaegle, Nikhil Parthasarathy, Relja Arandjelovic, João Carreira, and Olivier Hénaff. "Where should i spend my flops? efficiency evaluations of visual pre-training methods." arXiv preprint arXiv:2209.15589 (2022).

Thank You!!

Questions? Questions?

Appendix-A

RotNet Data-Prep Code Walkthrough

Each datapoint shape : (4 , C , H , W)

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

RotNet Data-Prep Code Walkthrough

Intial Setup

```python
class RotNetDataset(Dataset):
    """
    Adopts the RotNet data preparation.
    """

    def __init__(self, root, transform=None):
        self.transform = transform
        self.imgs = glob(root + "/**/*.jpg")

```

Each datapoint shape : (4 , C , H , W)

RotNet Data-Prep Code Walkthrough

Implementing the ***__getitem__()*** method

```python
def __getitem__(self, idx):
        img_path = self.imgs[idx]
        img = Image.open(img_path)
        label_0 = torch.LongTensor([0])
        label_1 = torch.LongTensor([1])
        label_2 = torch.LongTensor([2])
        label_3 = torch.LongTensor([3])
        img_0_tensor = self._img_prep(img, label_0)
        img_1_tensor = self._img_prep(img, label_1)
        img_2_tensor = self._img_prep(img, label_2)
        img_3_tensor = self._img_prep(img, label_3)

        labels = torch.cat((label_0, label_1, label_2, label_3), dim=0)
        imgs = torch.cat(
            (img_0_tensor, img_1_tensor, img_2_tensor, img_3_tensor), dim=0
        )

        return imgs, labels
```

Each datapoint shape : (4 , C , H , W)

RotNet Data-Prep Code Walkthrough

Each datapoint shape : (4 , C , H , W)
Implementing the augmentation routine.

```python
def _img_prep(self, img, label):
        rotation = int(label.item() * 90)
        img = rotate(img, rotation)
        aug = transforms.Compose([transforms.ToTensor(), transforms.Resize((28, 28))])

        return aug(img).unsqueeze(0)
```

RotNet Data-Prep Code Walkthrough

Each datapoint shape : (4 , C , H , W)

Implementing the ***__len__()*** method

```python
def __len__(self):
        return len(self.imgs)
```

RotNet Data-Prep Code Walkthrough

Each datapoint shape : (4 , C , H , W)

Fixing Batches with custom collate

```python
def custom_collate(batch):
    batch_imgs, batch_labels = default_collate(batch)
    batch_imgs = batch_imgs.reshape(
        batch_imgs.shape[0] * batch_imgs.shape[1],
        batch_imgs.shape[2],
        batch_imgs.shape[3],
        batch_imgs.shape[4],
    )
    return batch_imgs, batch_labels
```

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

FACULTY OF ENGINEERING

Appendix-B

CFN Data-Prep Code Walkthrough

CFN Data-Prep Code Walkthrough

Intial Setup

```python
class PuzzleDataset(Dataset):
    """
    Sets up the dataset instance for puzzle based SSL.
    """

    def __init__(self, root):
        self.imgs = glob(root + "/*.png")
        self.combinations = list(permutations(range(0, 9)))
        self.crop_shape = (225, 225)
        self.patch_shape = (75, 75, 3)
        self.final_patch_shape = (3, 64, 64)
```

CFN Data-Prep Code Walkthrough

Implementing the ***__getitem__()*** method

```python
def __getitem__(self, idx):
    img = Image.open(self.imgs[idx]).resize((256, 256))
    img = transforms.RandomCrop(self.crop_shape)(img)
    img = np.array(img)
    img_patches = patchify(img, (self.patch_shape), 75)
    img_patch_tensor = torch.zeros(9, *self.final_patch_shape)

    img_num = 0
    for i in range(img_patches.shape[0]):
        for j in range(img_patches.shape[1]):
            img_patch = transforms.ToTensor()(img_patches[i, j, 0])
            img_patch_tensor[img_num] = transforms.RandomCrop(
                (self.final_patch_shape[1], self.final_patch_shape[2])
            )(img_patch)
            img_num += 1

    label = int(random.randint(0, len(self.combinations)))
    combination = self.combinations[label]
    img_patch_tensor = img_patch_tensor[list(combination)]
    return img_patch_tensor, torch.LongTensor([label])
```

CFN Data-Prep Code Walkthrough

Implementing the ***__len__()*** method

```python
def __len__(self):
        return len(self.imgs)
```

Appendix-C

FCN Colorization Data-Prep Code Walkthrough

FCN Colorization Data-Prep Code Walkthrough

Intial Setup

```python
class ColorizationDataset(Dataset):
    """
    Implements the SSL Dataset setup for Colourful Image Colorization paper.
    """

    def __init__(self, root):
        self.imgs = glob(root + "/*.png")
        self.transform = transforms.Compose(
            [transforms.ToTensor(), transforms.Resize((128, 128))]
        )
```

FCN Colorization Data-Prep Code Walkthrough

Implementing the ***__getitem__()*** method

```python
def __getitem__(self, idx):
    img = Image.open(self.imgs[idx])
    inp_img = img.convert('L')
    return img, inp_img
```

FCN Colorization Data-Prep Code Walkthrough

Implementing the *__len__()* method

```python
def __len__(self):
    return len(self.imgs)
```

Appendix-D

SimCLR Data-prep Code Walkthrough

SimCLR Data-prep Code Walkthrough

Intial Setup

```python
class SimCLRDataset(Dataset):
    """
    Implements the SimCLR Dataset
    """

    def __init__(self, data_path, image_size):
        self.data = glob(data_path + "/*.jpg")

        self.aug = transforms.Compose(
            [
                transforms.RandomResizedCrop(size=image_size, scale=(0.08, 1)),
                transforms.RandomApply(
                    [transforms.ColorJitter(0.8, 0.8, 0.8, 0.2)], p=0.8
                ),
                transforms.RandomGrayscale(p=0.2),
                transforms.RandomApply(
                    [
                        transforms.GaussianBlur(
                            int(0.10 * image_size[0]), sigma=(0.1, 2.0)
                        )
                    ],
                    p=0.5,
                ),
                transforms.ToTensor(),
            ]
        )
```

SimCLR Data-prep Code Walkthrough

Implementing the ***__getitem__()*** method

```python
def __getitem__(self, idx):
        img = Image.open(self.data[idx])

        img_1 = self.aug(img)
        img_2 = self.aug(img)

        return img_1 , img_2
```

SimCLR Data-prep Code Walkthrough

Implementing the ***__len__()*** method

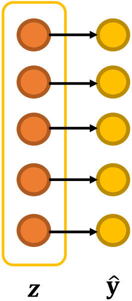
```python
def __len__(self):
    return len(self.data)
```

Appendix-E

Formulating SphereFace

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

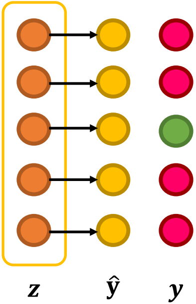
Formulating SphereFace



$$\hat{y} = \frac{exp\{z\}}{\sum_{k=1}^{K} exp\{z_k\}}$$

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

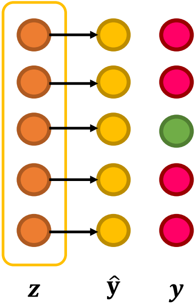
Formulating SphereFace



$$L = -\frac{1}{N}\sum_{i=1}^{N} y_i * \log\{\hat{y}_i\}$$

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

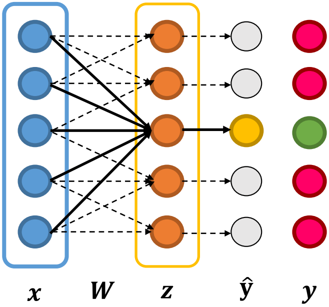
Formulating SphereFace



$$L = -\frac{1}{N}\sum_{i=1}^{N} y_i * \log\frac{exp\{z_i\}}{\sum_{k=1}^{K} exp\{z_{i,k}\}}$$

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

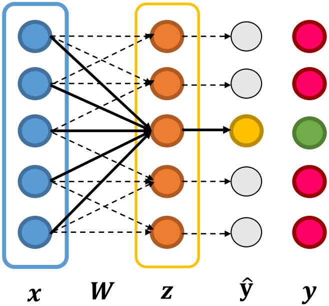
Formulating SphereFace



$$L = -\frac{1}{N}\sum_{i=1}^{N} y_i * \log \frac{exp\{W_i^T x_i\}}{\sum_{k=1}^{K} exp\{W_{i,k}^T x_i\}}$$

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

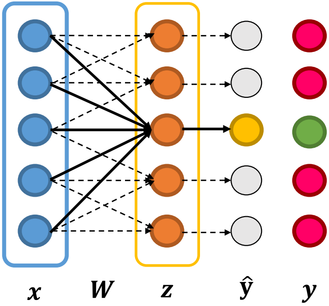
Formulating SphereFace



$$L = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{exp\{W_{y_i}^T x_i\}}{\sum_{k=1}^{K} exp\{W_{i,k}^T x_i\}}$$

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

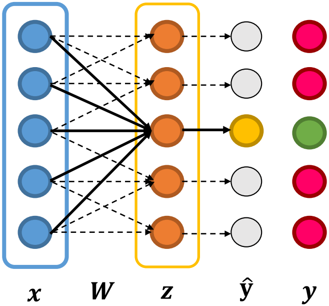
Formulating SphereFace



$$L = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{exp\{\left\|W_{y_i}\right\| \|x_i\| cos(\theta)\}}{\sum_{k=1}^{K} exp\{\|W_{i,k}\| \|x_i\| cos(\theta)\}}$$

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

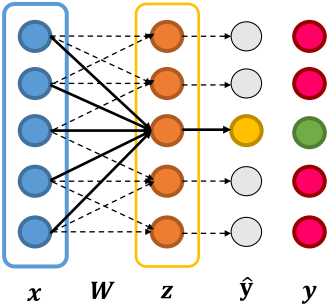
Formulating SphereFace



$$L = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{exp\{\|x_i\|cos(\theta)\}}{\sum_{k=1}^{K} exp\{\|x_i\|cos(\theta)\}}$$

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.

Formulating SphereFace



$$L = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{exp\{\|x_i\|\psi(\mu\theta)\}}{\sum_{k=1}^{K} exp\{\|x_i\|\psi(\mu\theta)\}}$$

Liu, Weiyang, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. "Sphereface: Deep hypersphere embedding for face recognition.", CVPR 2017.