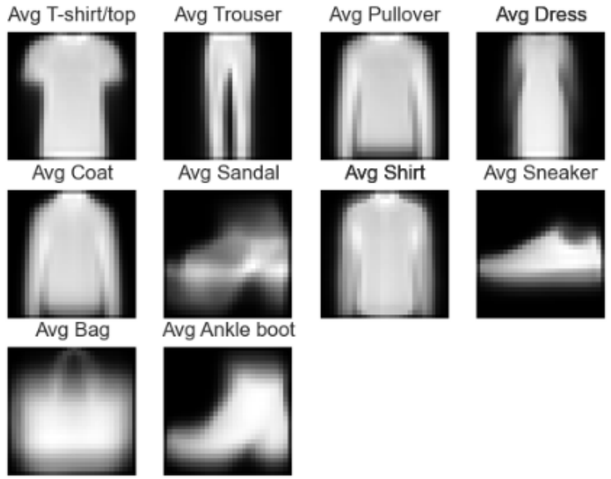
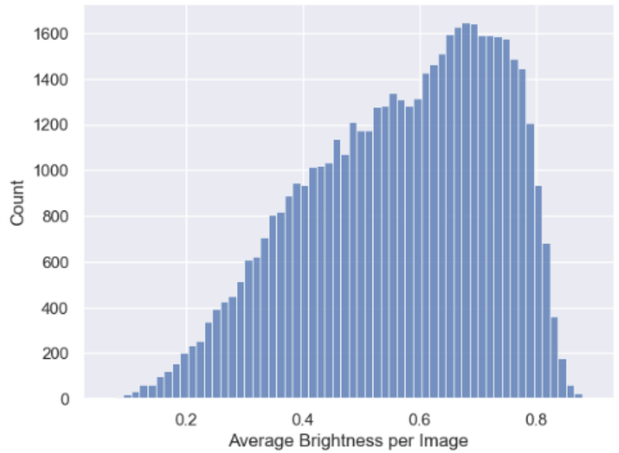
Fashion MNIST

Deep Learning - Part A

by Kenneth Chen





EDA: Basic

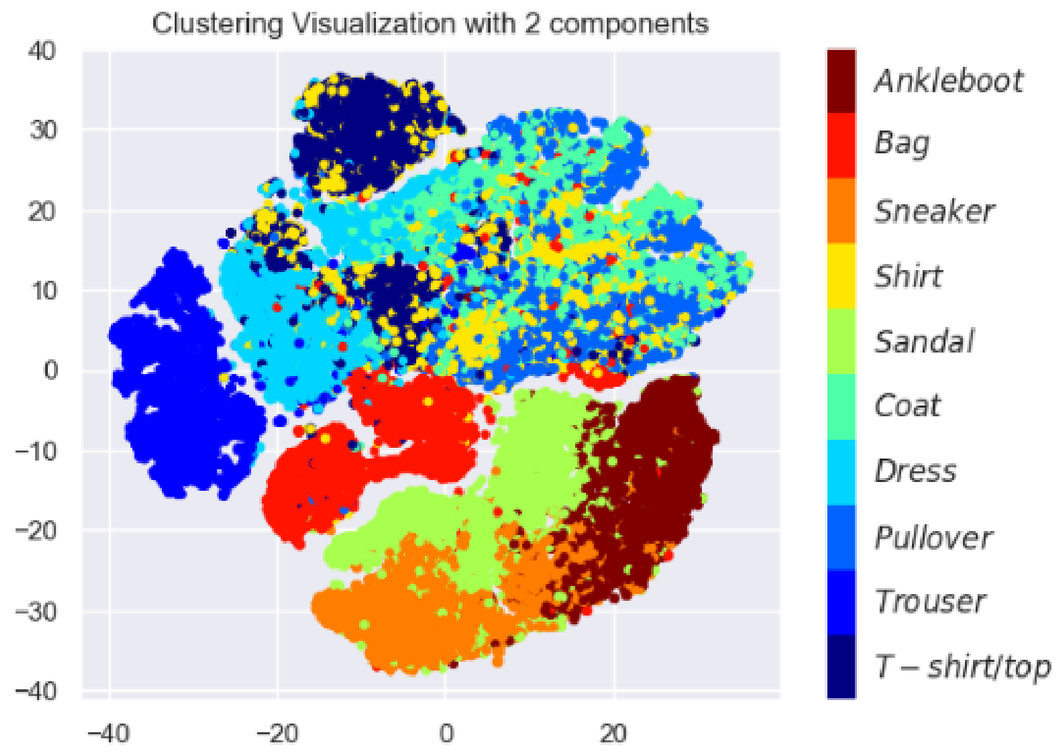
- Taking a look at the average image, items such as T-shirt, Trouser and Coat have **low variance** in terms of how different the images can vary

- Sandals, Bag, Boot and Dress seem to be much blurrier, which tells us there are **more unique designs of each**

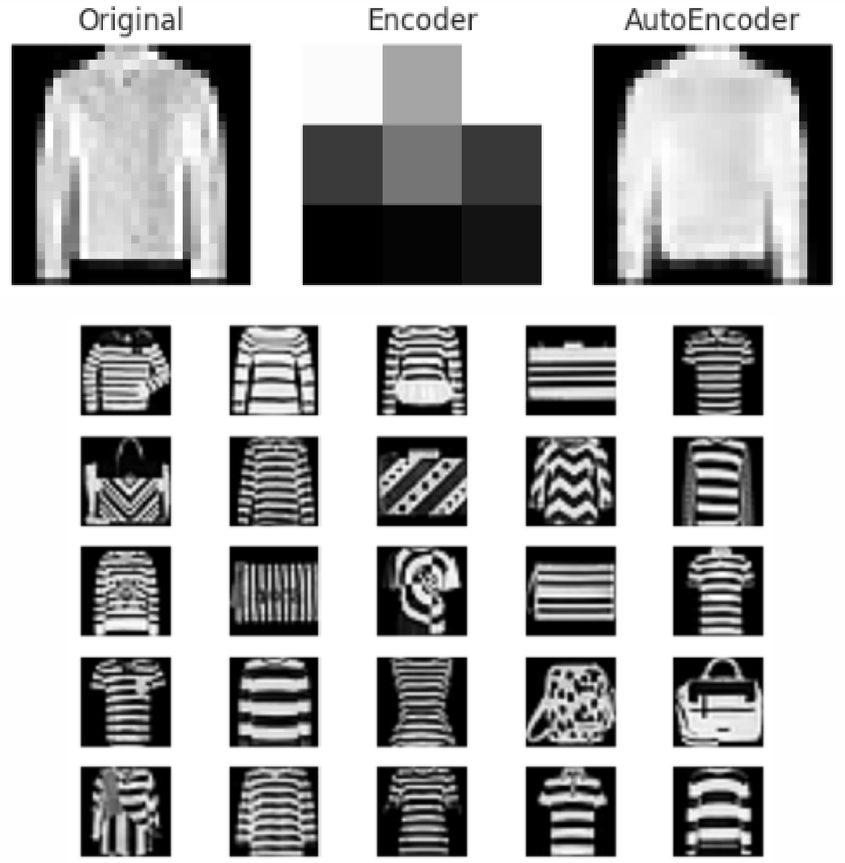
- The average brightness is skewed towards the left. Normalization may prove to be helpful.

EDA: Outlier Detection 😯

1. **There is potential for outliers as** dots spread all over.



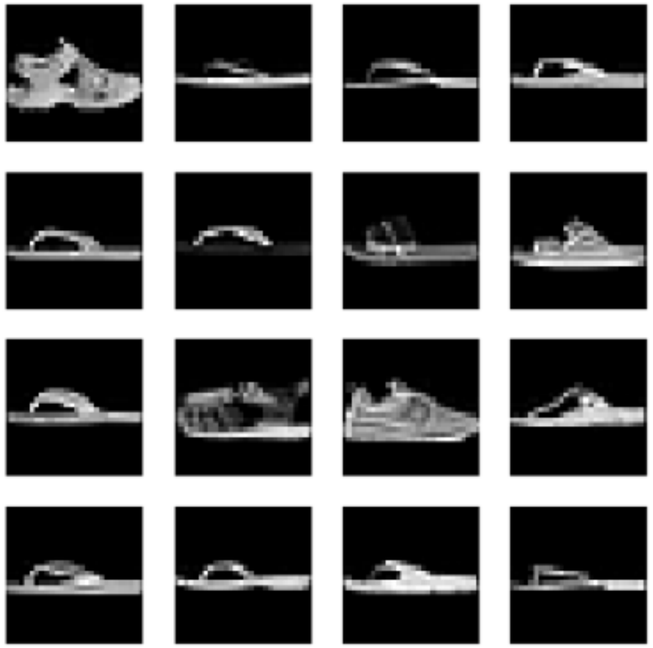
2. Trained a Conv Auto Encoder to **identify outliers**



EDA: Do all shoes point to the left?

Convolutions **do not** possess rotational equivariance. Augmentation without good reason will just **feed useless data** into our model. Do **we need to flip our data?**

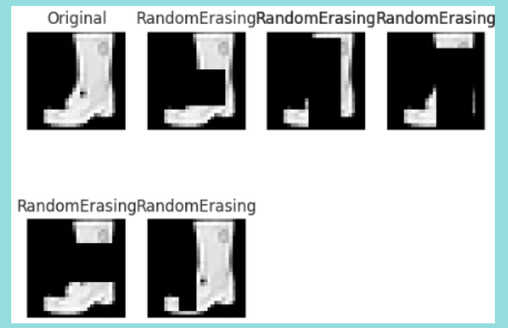
36 random samples of footwear



Developed and designed a custom **Slope Regression Algorithm** to identify right-pointing footwear

Modeling Process

Adam Optimizer + Aug:
Rotation, Flip and AugErase



1 — LENET

Val: 0.9012
Aug Val: 0.9054

Basic LeNet.

2 — RESNET18

Val: 0.9154
Aug Val: 0.9182

Standard ResNet18 architecture.

3 — VGG13

Val: 0.9321
Aug Val: 0.9384

VGG architecture tuned to fit the small image size of Fashion MNIST.

4 — MODIFIED MOBILENETV2

Val: 0.9314
Aug Val: 0.9297

Altered downsampling and stages for MBConv blocks.

5 — CUSTOM COATNET-0

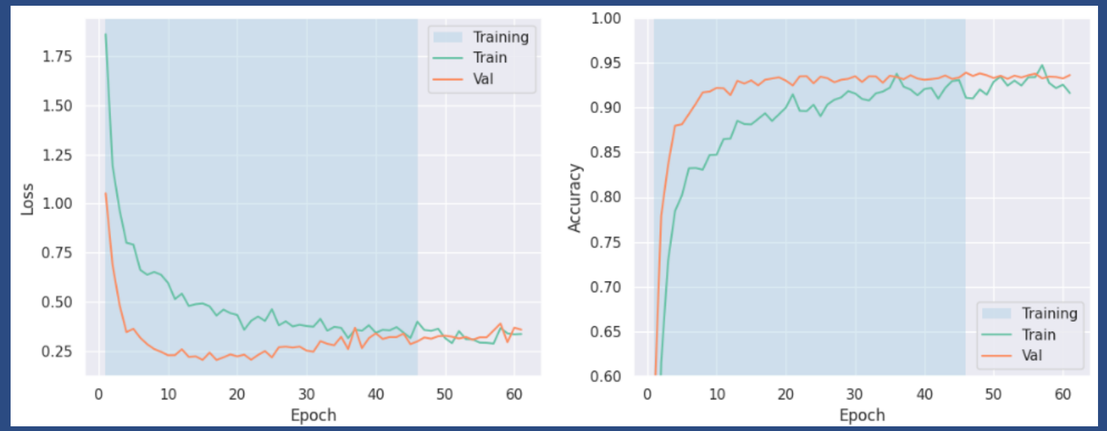
Val: 0.9320
Aug Val: 0.9361

Custom version of Convolution and Self Attention mechanism.

Testing CutMix augmentation on VGG13 and CoAtNet-0.

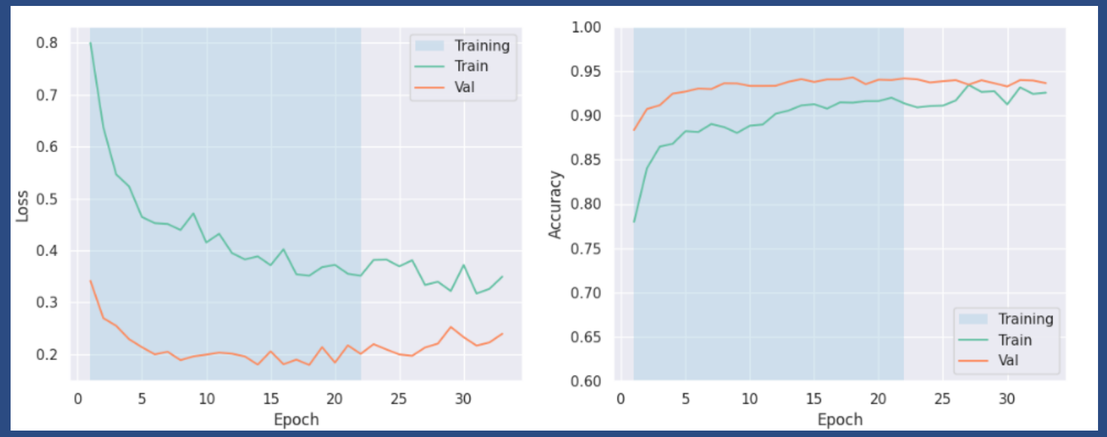
VGG13

Val: +0.001



Custom CoAtNet-0

Val: +0.005



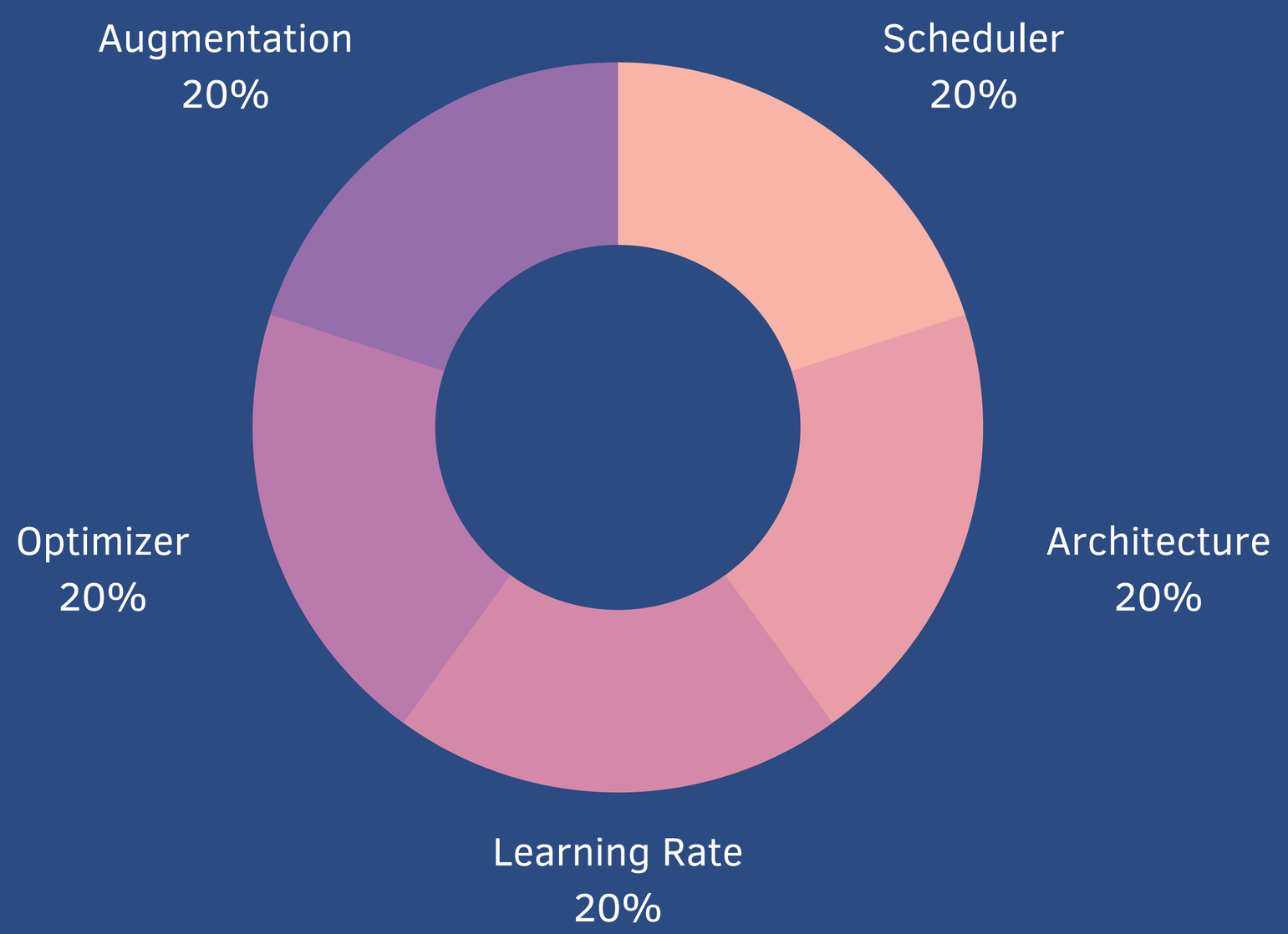
Model Improvement

Augmentation:
- CutMix
- AugErase

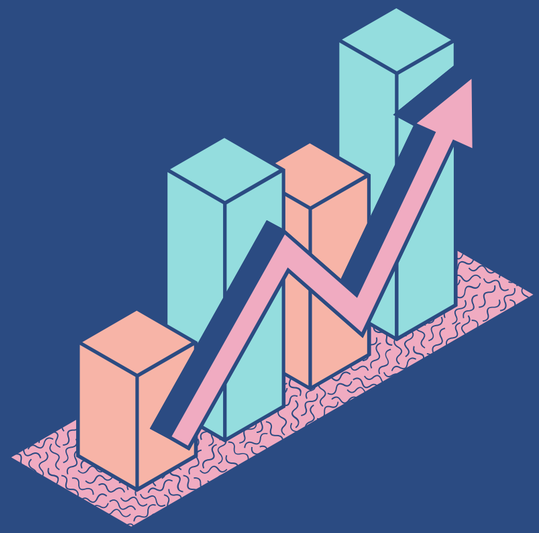
Optimizer:
- SGD
- AdamW
- Adam

Scheduler:
- CosineAnnealingLR
- StepLR

Architecture:
- CoAtNet-0
- GELU_VGG13
- VGG13



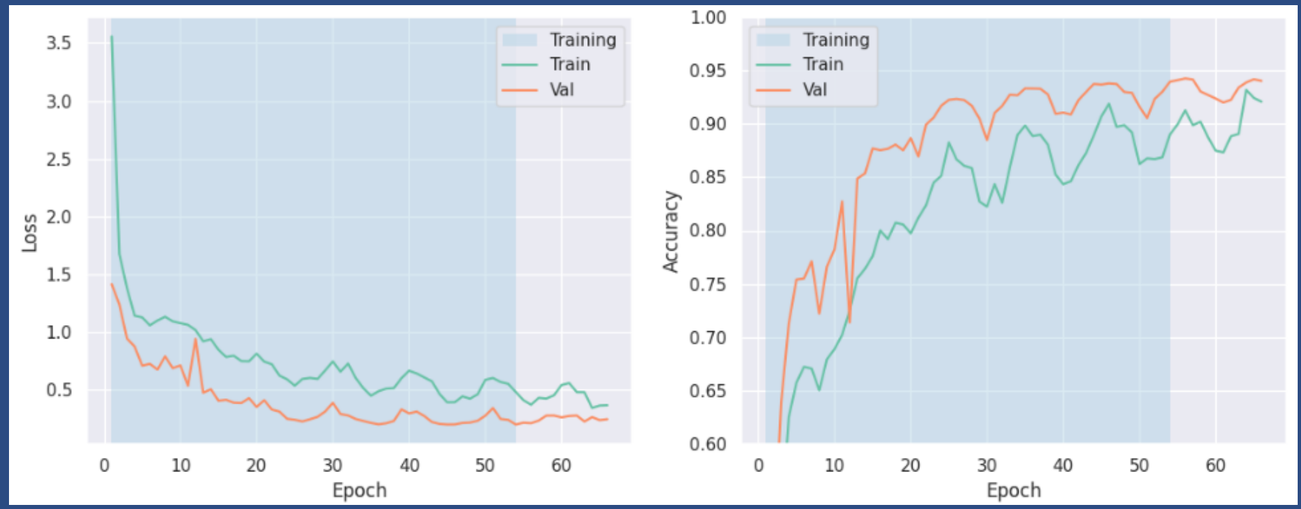
Final Model:



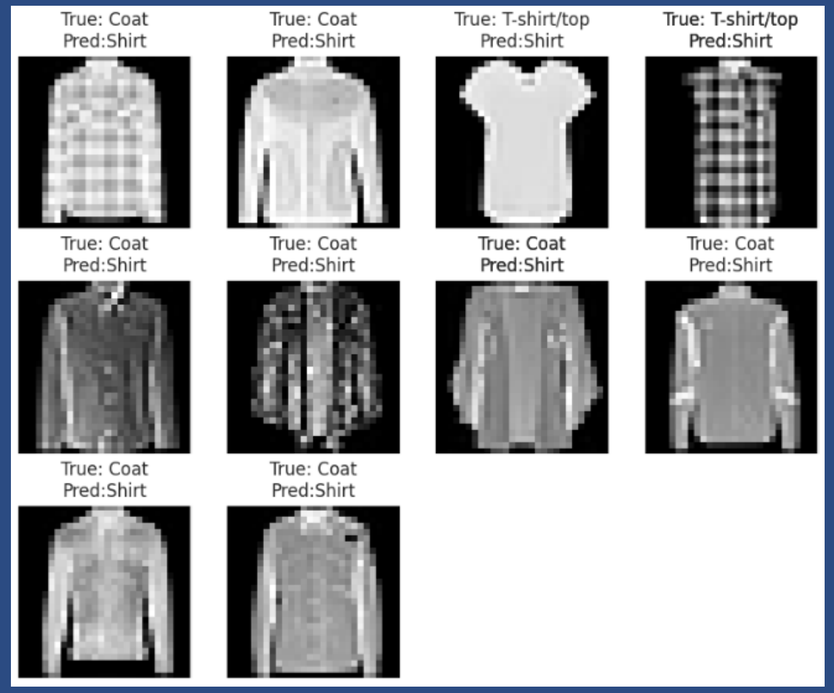
Augmentation: **AugErase + CutMix**

Model: **VGG13**
Optimizer: **AdamW**
Scheduler: **CosineAnnealingLR**

Train: 0.94
Val: 0.939
Test: 0.942



Error Analysis



Images with the worst CrossEntropyLoss

Struggle to determine the difference between Coats, Shits and T-shirt/top.

As seen previous by TSNE, we observe there to be a large overlap of features. The model struggles to identify the differences between them.

ITEM	ACCURACY
T-shirt/top	0.853
Pullover	0.880
Coat	0.887
Shirt	0.911
Dress	0.97
Sneaker	0.973
Sandal	0.979
Ankle Boot	0.985
Bag	0.994
Trouser	0.995

Error Analysis

The classwise accuracies also suggest the same.

We once again see this struggle by the model when it comes to Coats, Shirts, Pullover and T-shirt/top. Likely a deeper model may address this issue as downsampling could have occurred too fast.

Jump in accuracy from 'Shirt' to 'Dress'

This suggests that it is particularly the category of these upperbody apparels that has a harder to distinguish property.

Thank you

References

- Adam, A. (2018) Early Stopping and its Faults, Early Stopping and its Faults. Available at: https://alexadam.ca/2018/08/03/early-stopping/ (Accessed: 25 October 2022).
- Burnham, K.P. and Anderson, D.R. (2010) Model selection and multimodel inference: a practical information-theoretic approach. 2. ed. New York, NY: Springer.
- Dai, Z. et al. (2021) ‘CoAtNet: Marrying Convolution and Attention for All Data Sizes’. arXiv. Available at: http://arxiv.org/abs/2106.04803 (Accessed: 13 November 2022).
- He, K. et al. (2015) ‘Deep Residual Learning for Image Recognition’. arXiv. Available at: http://arxiv.org/abs/1512.03385 (Accessed: 13 November 2022).
- Howard, A.G. et al. (2017) ‘MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications’. arXiv. Available at: http://arxiv.org/abs/1704.04861 (Accessed: 13 November 2022).
- Hu, J. et al. (2019) ‘Squeeze-and-Excitation Networks’. arXiv. Available at: http://arxiv.org/abs/1709.01507 (Accessed: 13 November 2022).
- Shorten, C. and Khoshgoftaar, T.M. (2019) ‘A survey on Image Data Augmentation for Deep Learning’, Journal of Big Data, 6(1), p. 60. Available at: https://doi.org/10.1186/s40537-019-0197-0.
- Valueva, M.V. et al. (2020) ‘Application of the residue number system to reduce hardware costs of the convolutional neural network implementation’, Mathematics and Computers in Simulation, 177, pp. 232–243. Available at: https://doi.org/10.1016/j.matcom.2020.04.031.
- Xiao, H., Rasul, K. and Vollgraf, R. (2017) ‘Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms’. arXiv. Available at: http://arxiv.org/abs/1708.07747 (Accessed: 19 October 2022).
- Zhong, Z. et al. (2017) ‘Random Erasing Data Augmentation’. arXiv. Available at: http://arxiv.org/abs/1708.04896 (Accessed: 24 October 2022).