

FASHION IN THE CITY

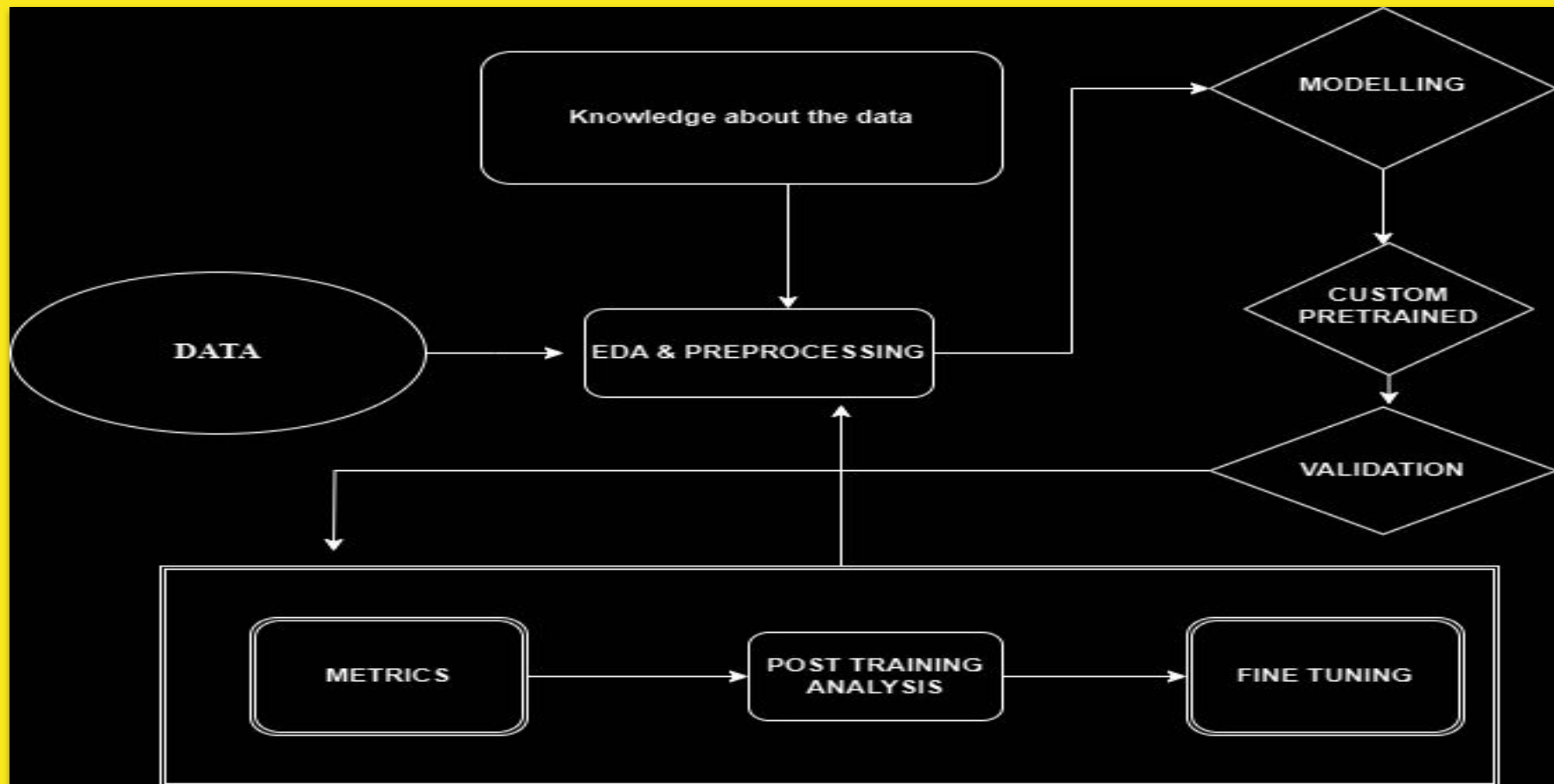
A computer vision project that takes image
input and outputs the attributes of the image

MOTIVATION

- Global fashion industry was estimated at **\$1.7 Trillion** in 2023
- United States, the fashion industry is valued at **\$343.70 Billion**
- **\$1460** per capita spending on clothing and footwear, the United States leads the pack as the top per capita spending on fashion
- **36%** of Gen Z buy new clothing at least every month
- Computer vision in healthcare, sports, automobile industries but not so much work has been done in the fashion industry
- personalized recommendable both affordable, reproducible







DATA

Image file and annotation folder with 331,213 samples of data.

- Image ID
- Encoded Pixels
- Height * Width
- ClassID

Json Data

- ID
- Name
- Supercategory
- 45 attributes



top, t-shirt, sweatshirt
pocket
neckline
pants



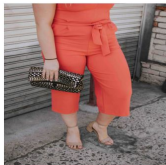
shirt, blouse
hat
tie
collar
lapel
sleeve
pocket
jacket
pants



shirt, blouse
belt
shoe
collar
lapel
sleeve
pocket
buckle
jacket
pants

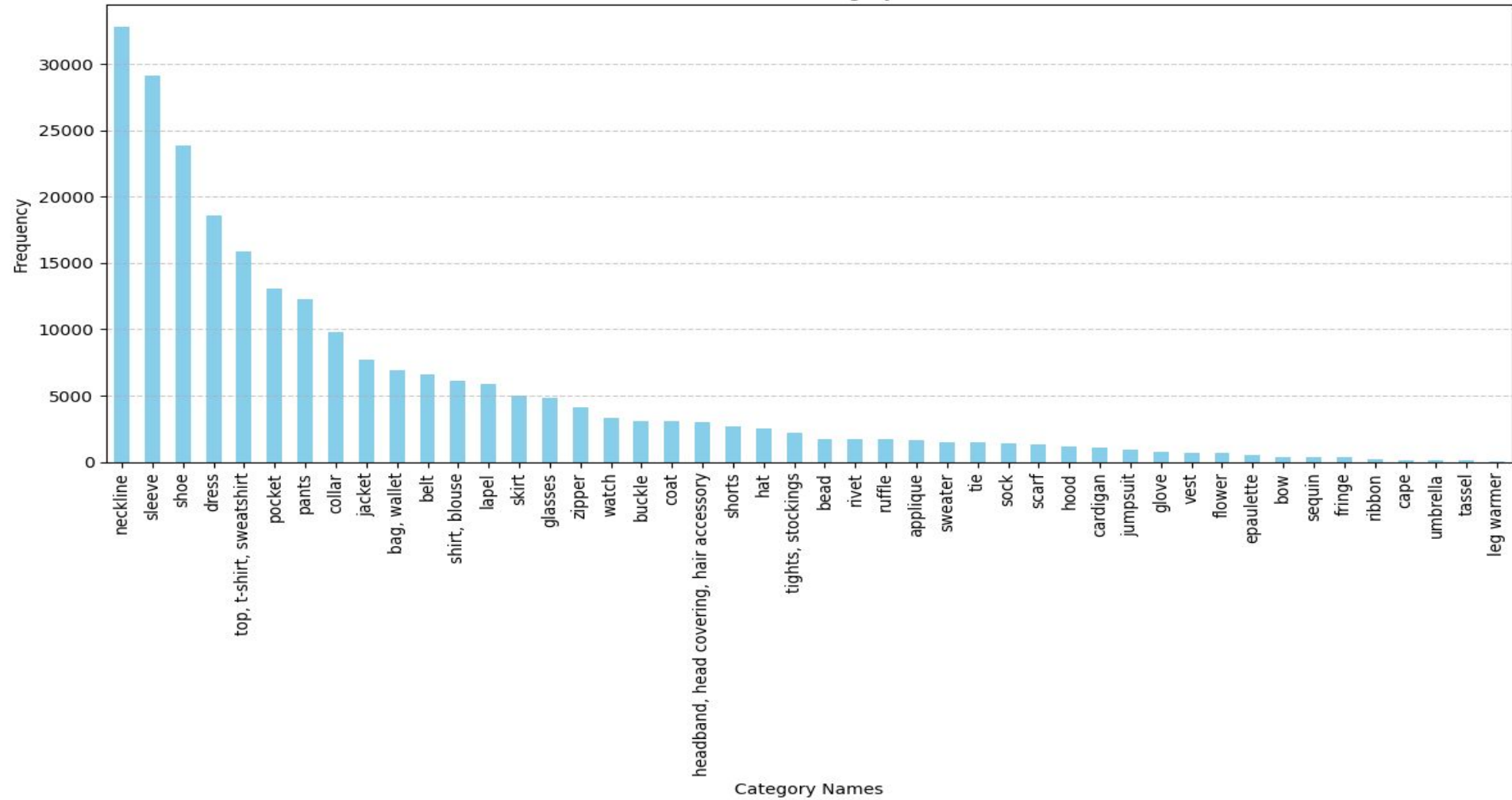


top, t-shirt, sweatshirt
sweater
sleeve
neckline

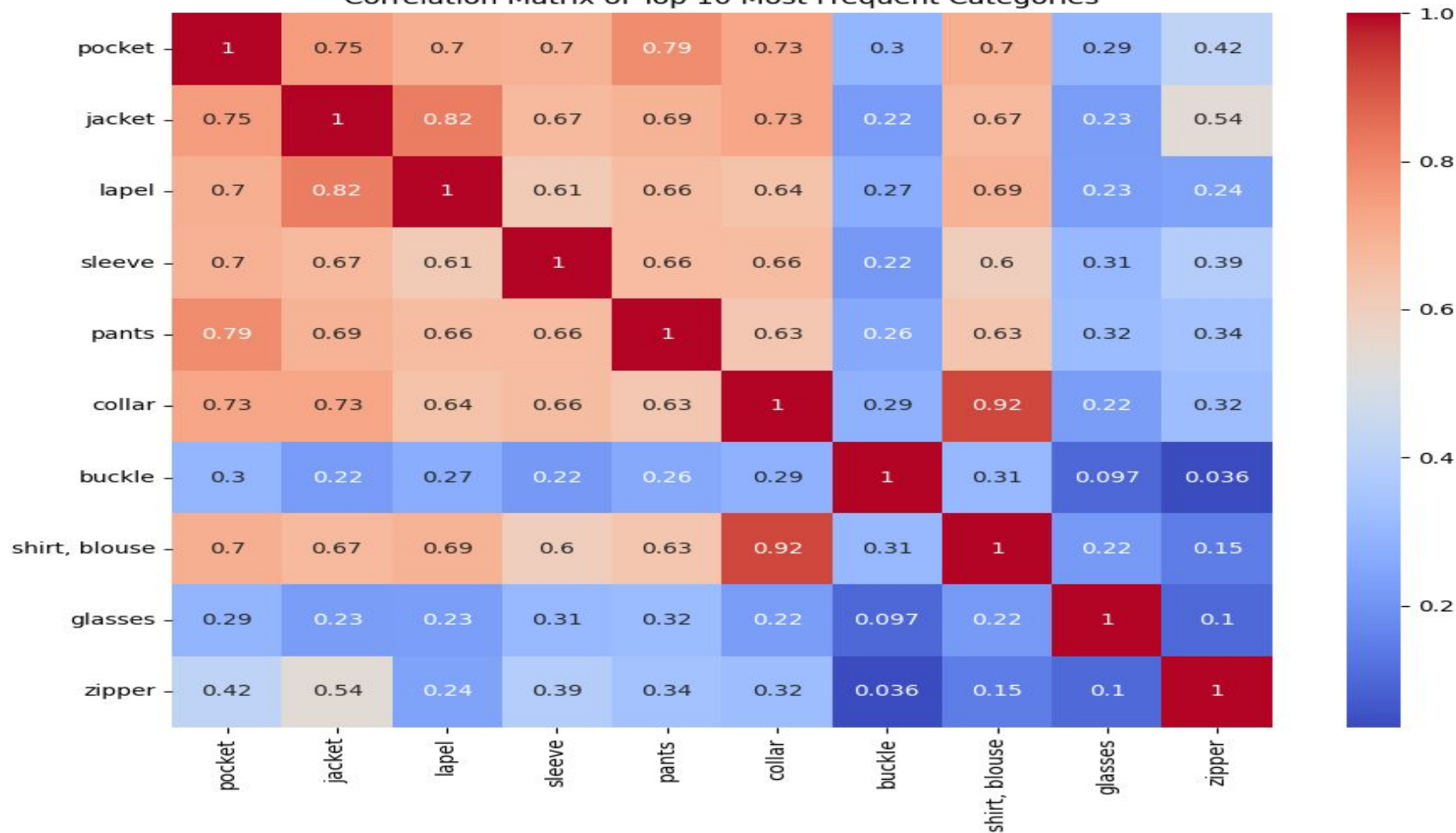


jumpsuit
belt
shoe
bag, wallet

Distribution of Category Names



Correlation Matrix of Top 10 Most Frequent Categories



$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n)$$

$$S(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i - m, j - n) K(m, n)$$

For a 3*3 kernel size the equation becomes,

$$S(i, j) = (K * I)(i, j) = \sum_{m=1}^3 \sum_{n=1}^3 I(i - m, j - n) K(m, n)$$

$$A(i, j) = \begin{cases} 0 & \text{if } S(i, j) < 0 \\ S(i, j) & \text{if } S(i, j) \geq 0 \end{cases}$$

MODELS

- Resnet50
- Inceptionnet
- ViT B 16
- Fasterrcnn_resnet50
- Resnet 101
- CNN

METRIC

	Accuracy	F1 Score	Precision	Recall
ViT_B_16	0.150348	0.748433	0.784978	0.715139
ResNet_101	0.074455	0.650119	0.690595	0.614126
CNN	0.037504	0.538761	0.67368	0.448866
Inceptionnet	0.02	0.52	0.65	0.34

POST TRAINING ANALYSIS

- Assign class weights to address multilabel imbalance
- FocalLoss loss function
- Regularization (Dropout and residual skip connection)
- Alternate between max pooling and average pooling

DEMONSTRATION

FUTURE WORK AND APPLICATION

- Committee of network
- Add text to image implementation
- Model can identify additional attributes such as color, design etc
- Image segmentation
- Fashion designers bespoke designs for customers