Fashion Classification

By: Abhinav Sharma, Edgar Leon, Shazia Nooruddin

Berkeley
UNIVERSITY OF CALIFORNIA

Fashion Classification Problem

- Task: Classify image data for accurate product representation
- Classes
  - Shirts (1322 images, 27%)
  - Pants (1139 images, 24%)
  - Shoes (1116 images, 23%)
  - Watches (1249 images, 26%)
- Characteristics
  - White background
  - Balanced
  - Variety of colors/patterns
  - Objects of interest are in the center
  - 2-5 KB
  - 60 x 80 pixels
  - Resolution: 96 dpi

Berkeley
UNIVERSITY OF CALIFORNIA

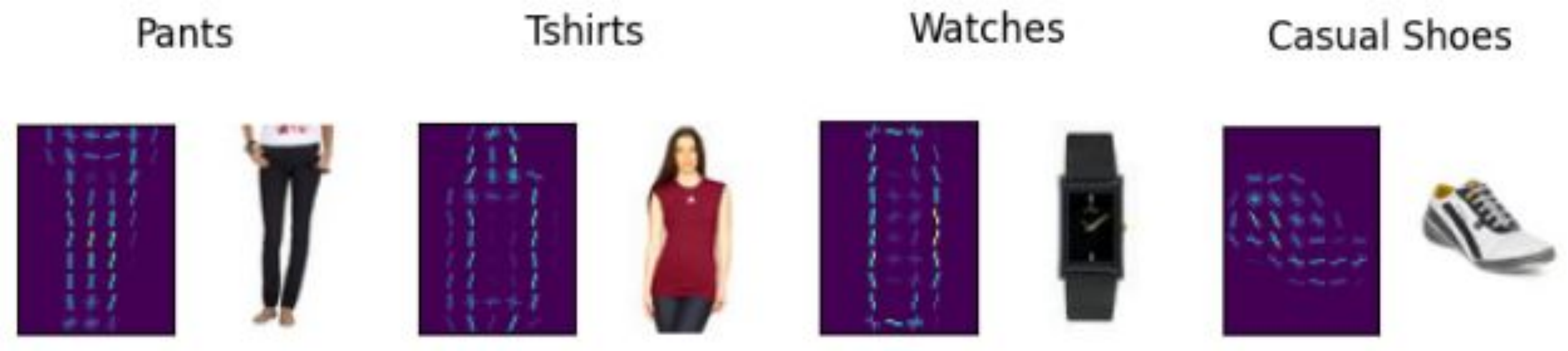
Preprocessing

- Resized images
- Images already centered
- PCA or UMAP transformation
- Train, test, validation (70%, 20%, 10%)
- Explored grayscale vs 3-channel RGB images

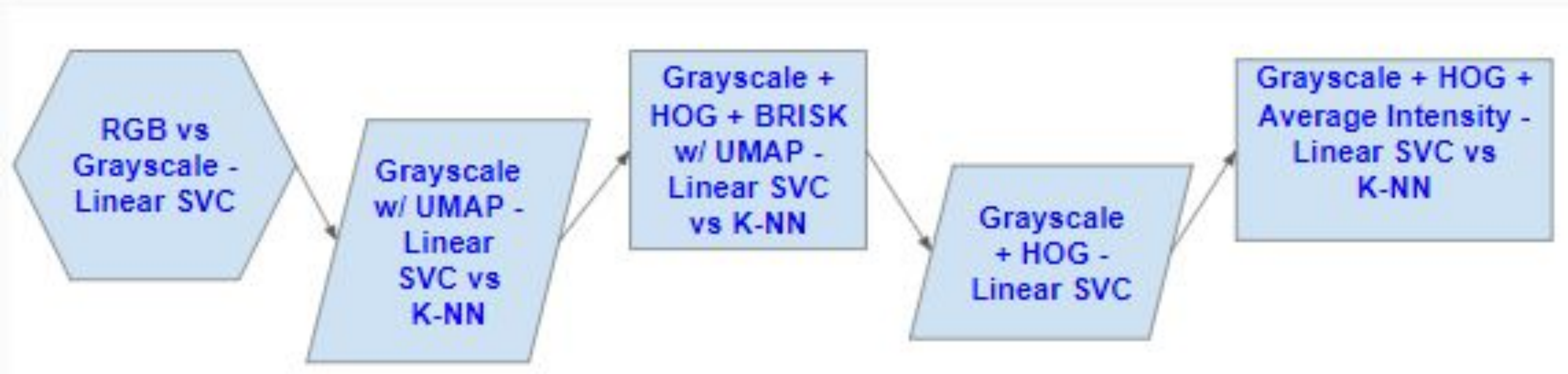
Feature Selection - Prelim Logit

Feature	Accuracy
HOG*	97.1%
HSV (Value) Histogram	45.3%
Harris Corner	55.7%
SIFT with BOVW	68.8%
BRISK*	47%
RGB Histogram	51%
ResNet101*	79%

HOG - Best Feature



Classification



RGB vs Grayscale - Linear SVC

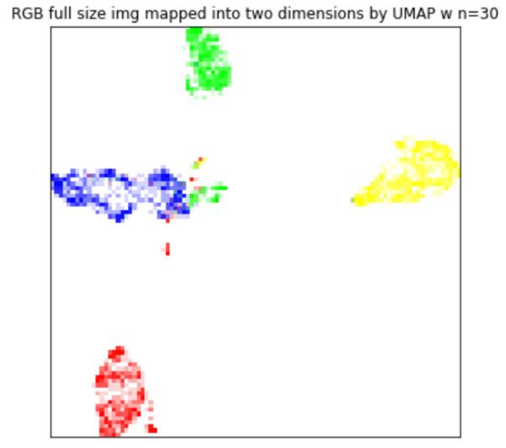
Grayscale w/ UMAP - Linear SVC vs K-NN

Grayscale + HOG + BRISK w/ UMAP - Linear SVC vs K-NN

Grayscale + HOG - Linear SVC

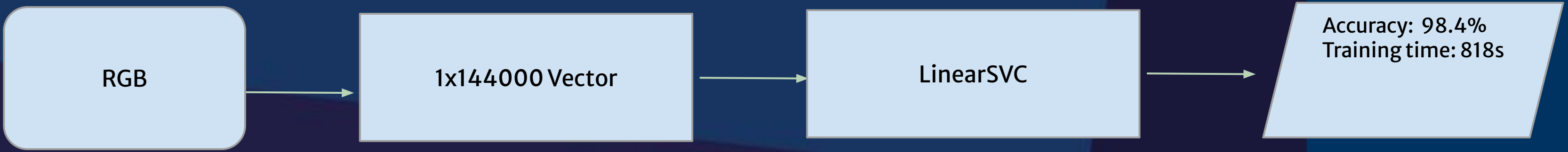
Grayscale + HOG + Average Intensity - Linear SVC vs K-NN

Classification RGB Image



RGB full size img mapped into two dimensions by UMAP w n=30

UMAP mapping of RGB train images to 2 dimensions



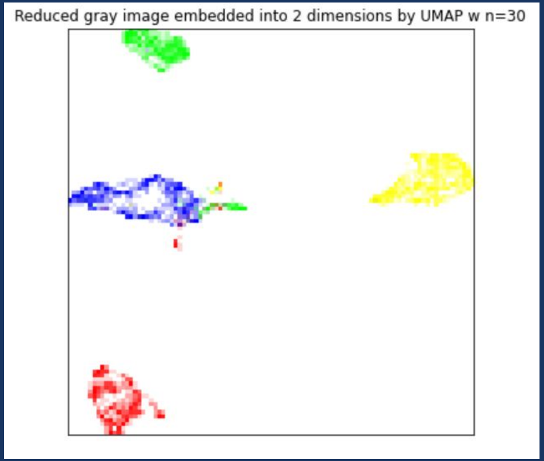
RGB

1x144000 Vector

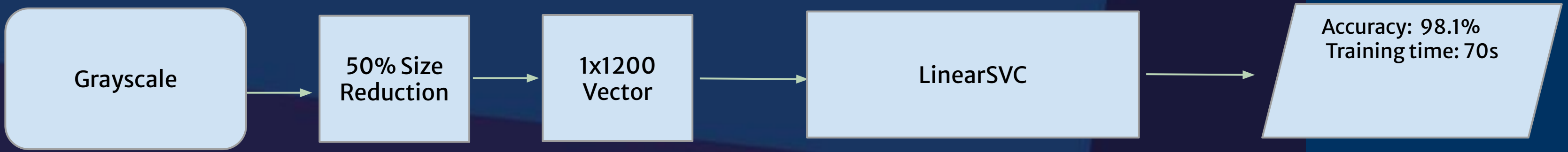
LinearSVC

Accuracy: 98.4%
Training time: 818s

Classification Grayscale img



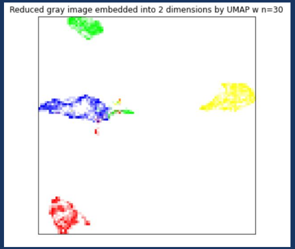
UMAP mapping of Grayscale train images to 2 dimensions



Accuracy: 98.1%
Training time: 70s

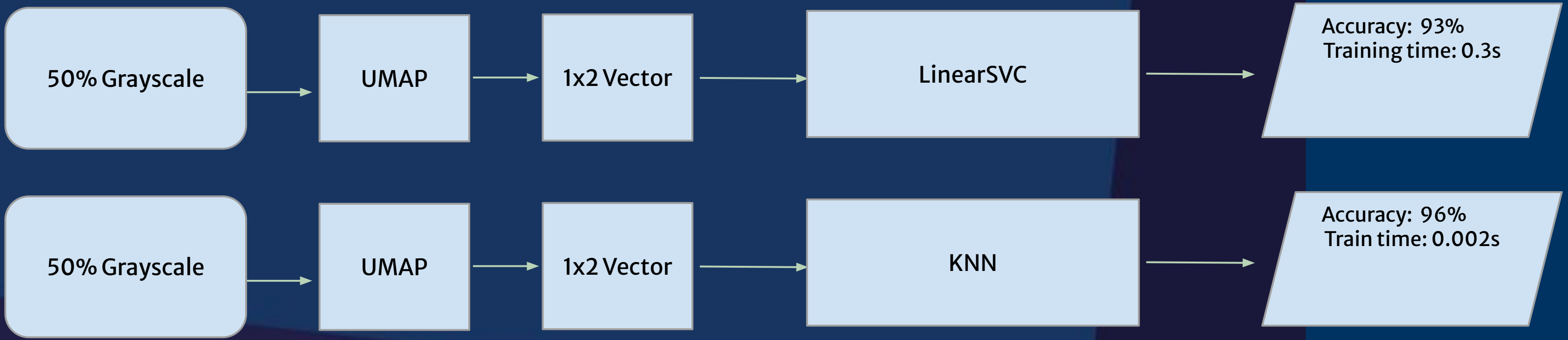
Moving forward we use reduced Grayscale images →

Classification



Reduced gray image embedded into 2 dimensions by UMAP w n=30

What if we could use the UMAP mapping for the training
Output is size 2 vector
UMAP learned architecture on Train data alone – Save it
Load saved model, transform Test data on UMAP model



50% Grayscale → UMAP → 1x2 Vector → LinearSVC →

Accuracy: 93%
Training time: 0.3s

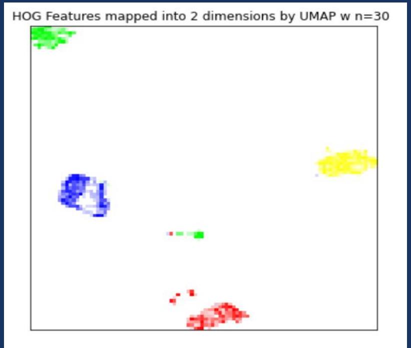
50% Grayscale → UMAP → 1x2 Vector → KNN →

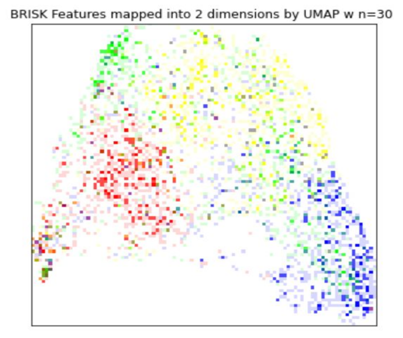
Accuracy: 96%
Train time: 0.002s

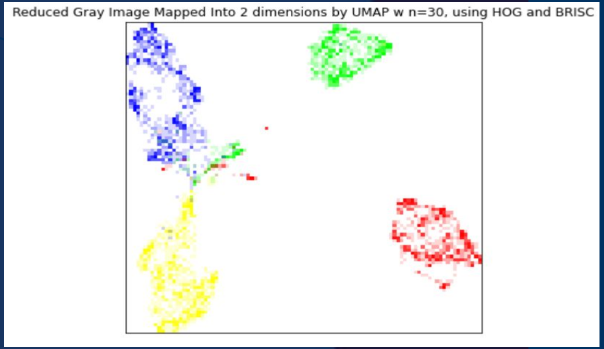
Want to use K Classifier and UMAP →only need accuracy to go up

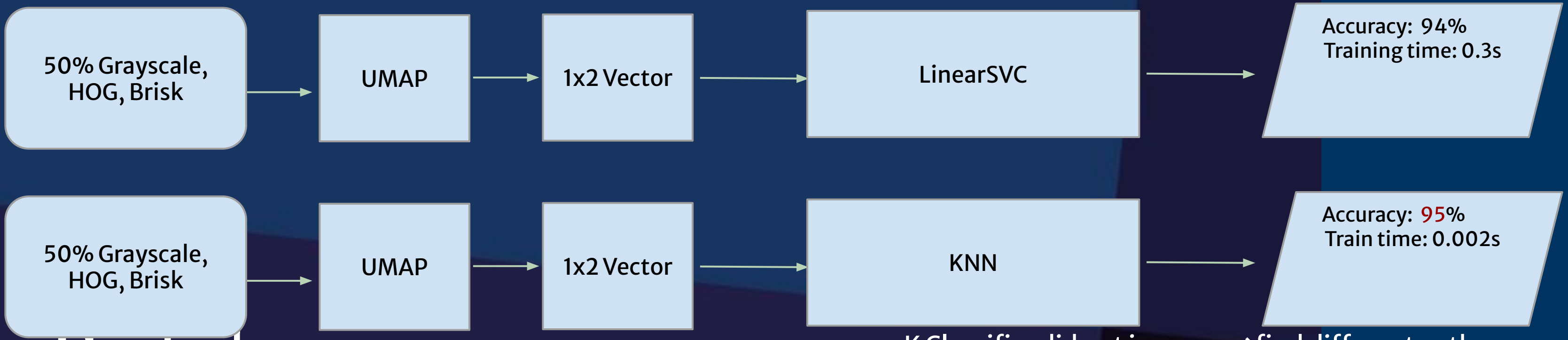
Berkeley
UNIVERSITY OF CALIFORNIA

UMAP Visualizations + Classifier


HOG Features mapped into 2 dimensions by UMAP w n=30


BRISK Features mapped into 2 dimensions by UMAP w n=30


Reduced Gray Image Mapped Into 2 dimensions by UMAP w n=30, using HOG and BRISC


50% Grayscale, HOG, Brisk

UMAP

1x2 Vector

LinearSVC

Accuracy: 94%
Training time: 0.3s

50% Grayscale, HOG, Brisk

UMAP

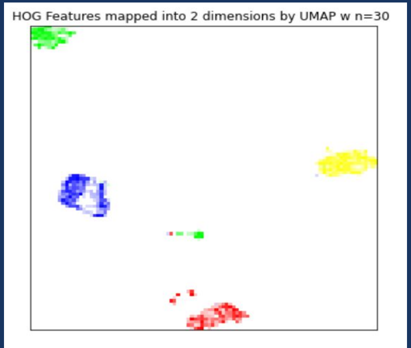
1x2 Vector

KNN

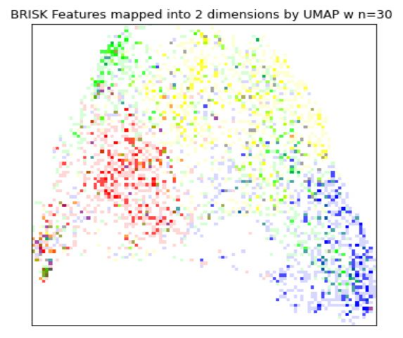
Accuracy: 95%
Train time: 0.002s

K Classifier did not improve →find different path

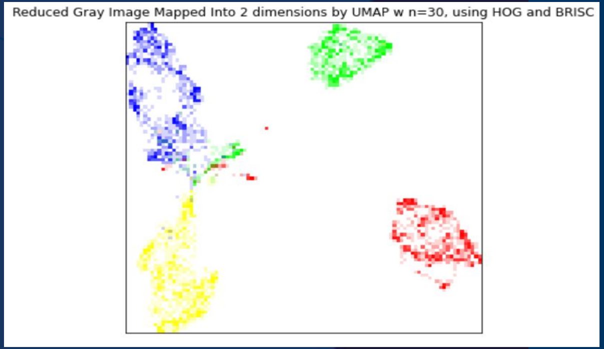
UMAP + Features



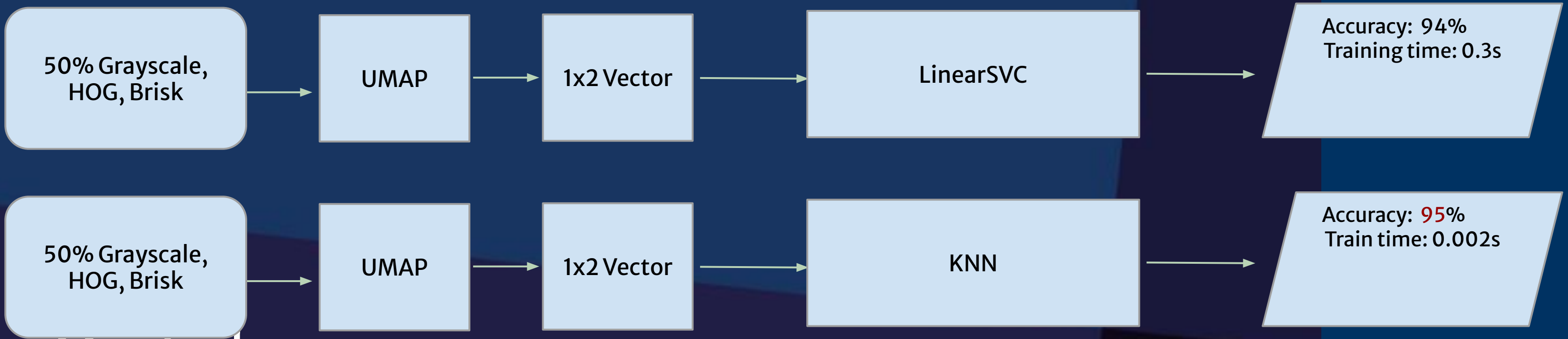
HOG Features mapped into 2 dimensions by UMAP w n=30



BRISK Features mapped into 2 dimensions by UMAP w n=30



Reduced Gray Image Mapped Into 2 dimensions by UMAP w n=30, using HOG and BRISC



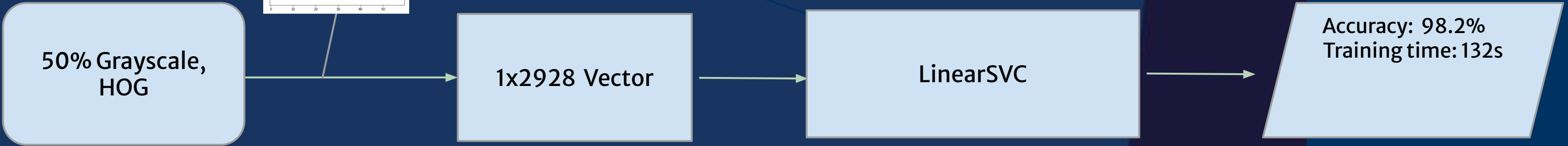
50% Grayscale, HOG, Brisk → UMAP → 1x2 Vector → LinearSVC → Accuracy: 94% Training time: 0.3s

50% Grayscale, HOG, Brisk → UMAP → 1x2 Vector → KNN → Accuracy: 95% Train time: 0.002s

Berkeley
UNIVERSITY OF CALIFORNIA

Gray and Hog

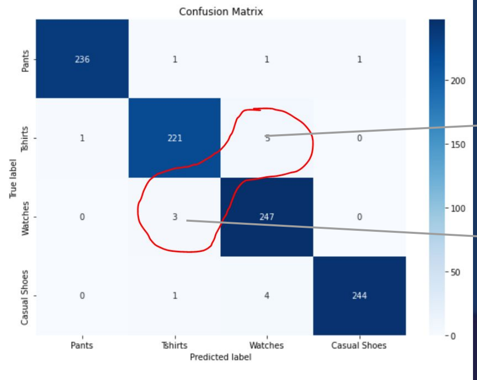
Item of Class: 'Tshirts' Labeled As 'Watches'


50% Grayscale, HOG

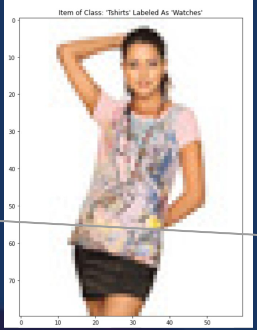
1x2928 Vector

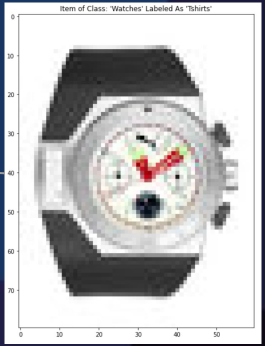
LinearSVC

Accuracy: 98.2%
Training time: 132s

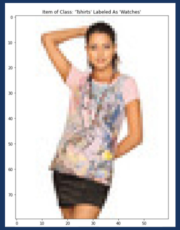

Confusion Matrix

True label \ Predicted label	Pants	Tshirts	Watches	Casual Shoes
Pants	236	1	1	1
Tshirts	1	221	5	0
Watches	0	3	247	0
Casual Shoes	0	1	4	244

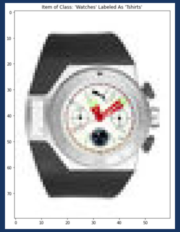

Item of Class: 'Tshirts' Labeled As 'Watches'


Item of Class: 'Watches' Labeled As 'Tshirts'

Gray and Hog



Item of Class: 'Tshirts' Labeled As 'Watches'

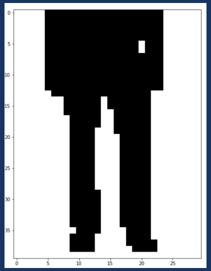


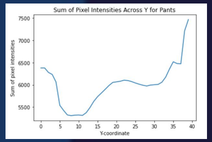
Item of Class: 'Watches' Labeled As 'Tshirts'

Considered face det, but..

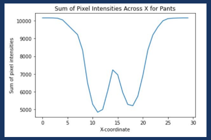
Top edge differentiates some Tshirts from some Watches.

Not being taken account of.



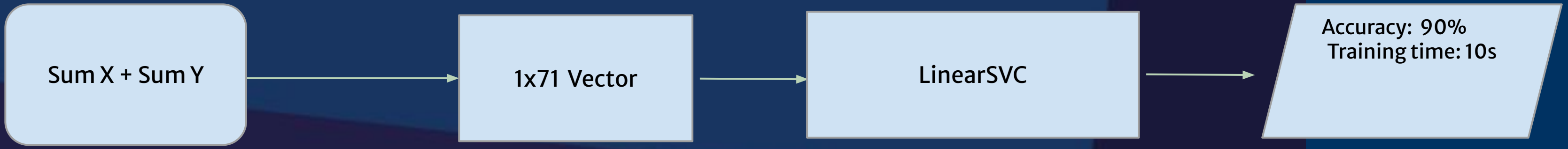


Sum of Pixel Intensities Across Y for Pants



Sum of Pixel Intensities Across X for Pants

Turned images to b&w
Sum of Y vector
Sum of X vector



Sum X + Sum Y

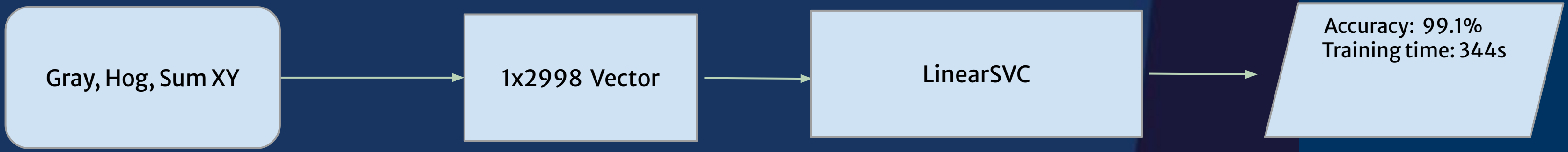
1x71 Vector

LinearSVC

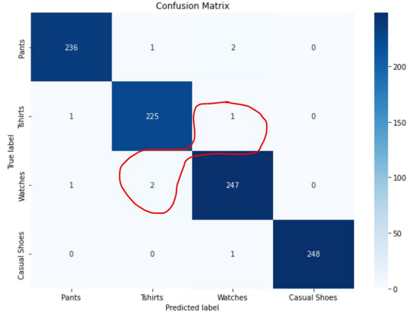
Accuracy: 90%
Training time: 10s

Berkeley
UNIVERSITY OF CALIFORNIA

Gray and Hog and Axis Intensity

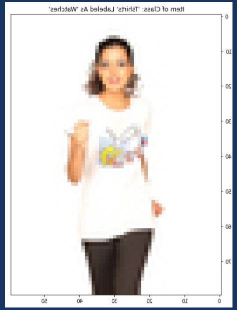


Gray, Hog, Sum XY → 1x2998 Vector → LinearSVC → Accuracy: 99.1% Training time: 344s

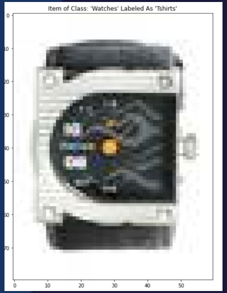


Confusion Matrix

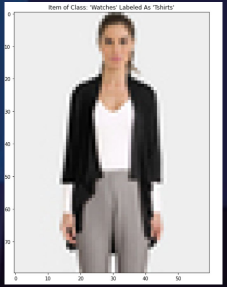
True label \ Predicted label	Pants	Tshirts	Watches	Casual Shoes
Pants	236	1	2	0
Tshirts	1	225	1	0
Watches	1	2	247	0
Casual Shoes	0	0	1	248



Fails because of white Tshirt



Failure unknown



Mislabeled picture.

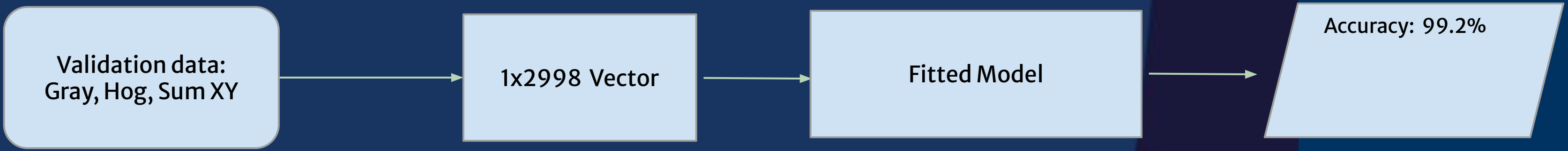
Labeled as Watch

Classifier identified as Tshirt

This is our best model →

Berkeley
UNIVERSITY OF CALIFORNIA

Model on Validation

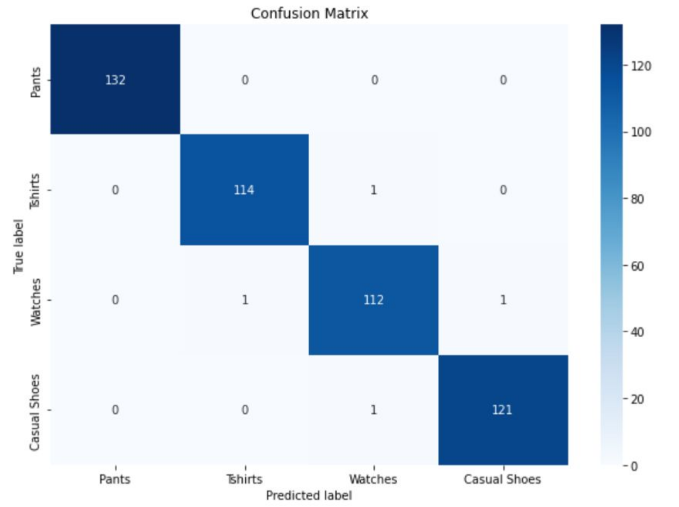


Validation data: Gray, Hog, Sum XY

1x2998 Vector

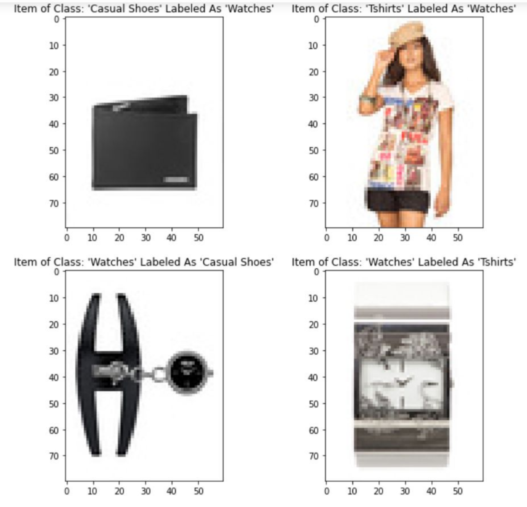
Fitted Model

Accuracy: 99.2%



Confusion Matrix

True label \ Predicted label	Pants	Tshirts	Watches	Casual Shoes
Pants	132	0	0	0
Tshirts	0	114	1	0
Watches	0	1	112	1
Casual Shoes	0	0	1	121



Item of Class: 'Casual Shoes' Labeled As 'Watches'
Item of Class: 'Tshirts' Labeled As 'Watches'
Item of Class: 'Watches' Labeled As 'Casual Shoes'
Item of Class: 'Watches' Labeled As 'Tshirts'
0
10
20
30
40
50
60
70

Classification Approach

Grayscale	UMAP	Features	Model	Accuracy
No	No	None	Linear SVC	98%
Yes	No	None	Linear SVC	98%
Yes	Yes	None	Linear SVC	93%
Yes	Yes	None	K-NN	96%
Yes	Yes	HOG, BRISK	Linear SVC	94%
Yes	Yes	HOG, BRISK	K-NN	95%
Yes	No	HOG	Linear SVC	98%
Yes	No	HOG, Avg Intensity	Linear SVC	99%
Yes	Yes	HOG, Avg Intensity	K-NN	25%

Classification Approach - Hyperparameter Tuning w/ GridSearchCV

From C = 0.00001 to 100000:

Grayscale	UMAP	Features	Model	Accuracy	Best Hyperparameters	Training Time	Prediction Time
Yes	No	HOG, Axis Intensity	Linear SVC	99.1%	C=1xe-5	6.73 s	0.025 s

The hyperparameter C in a LinearSVC model controls the trade-off between the margin size and the misclassification error. A larger C means that the model will try to find a smaller margin, but it may also overfit the training data and perform poorly on new data.

Berkeley
UNIVERSITY OF CALIFORNIA

Conclusion

Grayscale	UMAP	Features	Model	Accuracy	Hyperparameter	Training Time	Prediction Time
Yes	No	HOG, Axis Intensity	Linear SVC	99.1%	C=1xe-5	6.73 s	0.025 s

RGB	UMAP	Features	Model	Accuracy	Hyperparameter	Training Time	Prediction Time
Yes	No	RGB	Linear SVC	98%	C=1xe-5	121 s	0.82 s

Conclusion

- Increased accuracy
- Reduced file size
- Improved performance
- Intuitive Approach
- For future experiment more with UMAP

Berkeley
UNIVERSITY OF CALIFORNIA

Appendix: Class. Approach 2

- Collab/Google computing
- Hyperparameter tuning with PCA Transformed HOG and RestNet101 with models Logistic Regression and Linear SVC

Features	Model	Accuracy	Best Hyperparameters	Training Time	Prediction Time
HOG	Logistic Regression	98%	C=100, Penalty=L1, Solver=Saga	4.62 s	0.016 s
HOG	Linear SVC	98%	C=10, Gamma=0.01, Kernel=RBF	0.26 s	0.23 s
HOG + ResNet101	Linear SVC	97.8%	C=10, Gamma=auto, Kernel=RBF	0.21 s	0.11 s