THE UNIVERSITY OF CHICAGO

Shoe Image Classification

with.
Mia Song

Table of Content

Problem Statement

The Beginner's Guide to Selling on Amazon

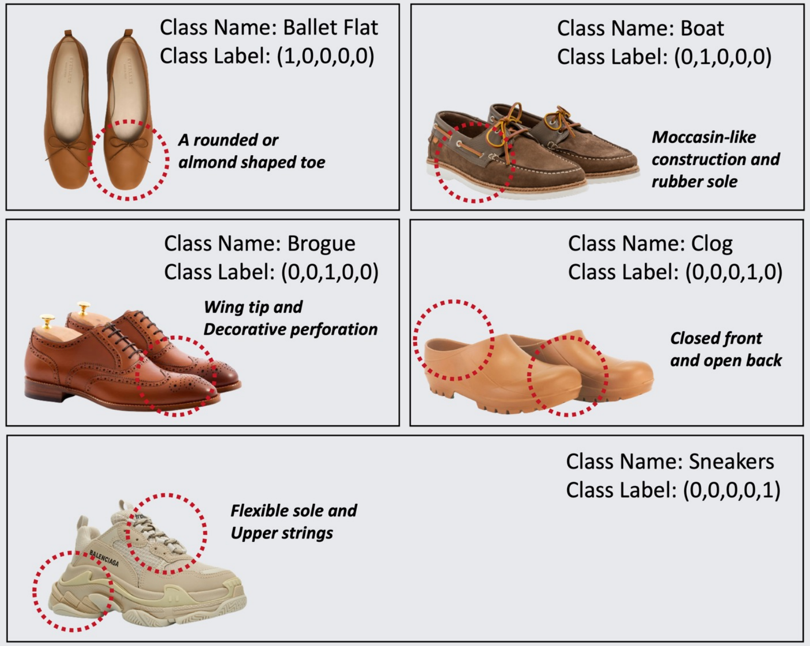
What you need to start listing products: In most cases, products must have a Global Trade Item Number (GTIN), such as a UPC, an ISBN, or an EAN. Amazon uses these product IDs to identify the exact item you're selling··· In addition to a product ID, here's some of the important information that goes into each product listing: Product title, Search terms, and relevant keywords···

- Product categorization is currently a manual, labor-intensive task.
  - This process is not only time-consuming and costly, but it is also subject to inaccuracies arising from subjective interpretation of product images and human error. These inaccuracies can subsequently impact customer satisfaction and sales negatively.
- We aim to develop a deep neural network model for product image classification, initially focusing on shoe images.
  - By automating the categorization task, we anticipate significant gains in efficiency and accuracy, as well as substantial cost reductions, revolutionizing our current process in the fashion industry.

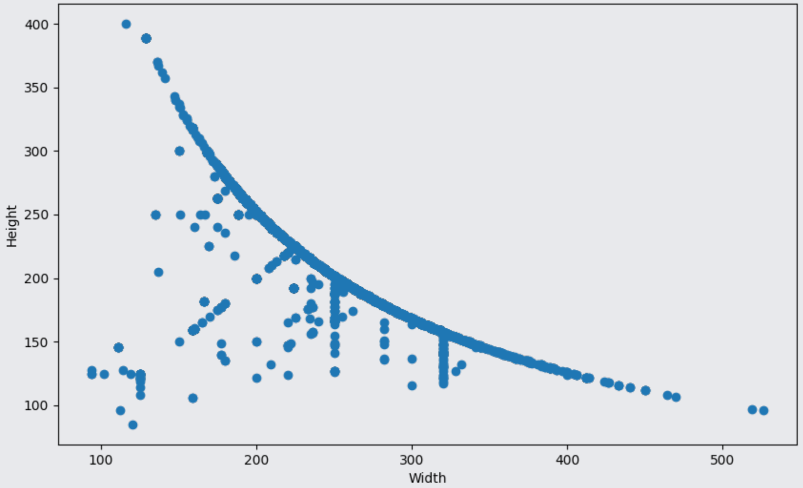
Exploratory Data Analysis

1. Class Distribution
   - Train Set: 10,000 images (2,000 per class)
   - Test Set: 1,215 images (Variable distribution across classes)
   - Validation Set: 2,500 images (500 per class)
2. Pixel Value Distribution: 0 ~ 255 pixels
3. Size Distribution: As depicted in the chart
4. Data Quality: No improper data types or low-resolution images were detected.

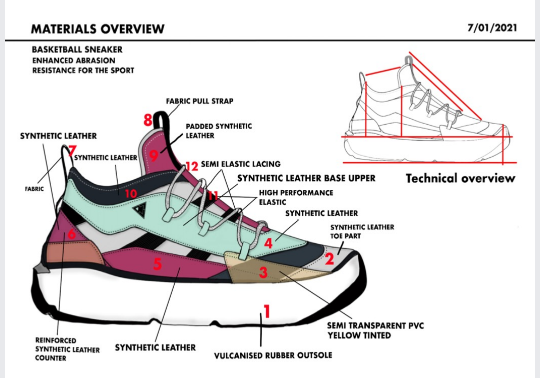
Class Information



Size Distribution (X: Width, Y: Height)



Data Assumptions & Hypotheses



Feature Complexity

Shoe designs exhibit complex features, including diverse designs, textures, materials, and patterns.

Inter-Class Variability

There is high inter-class variability, as different shoe types possess distinct features.

Scale and Position Invariance

Images consistently show shoes at a similar scale and orientation, with no hierarchical structure.

High Data Quality

The dataset is of high quality, with no dodgy or inappropriate images found upon review.

Candidate Models:
Convolutional Neural Networks (CNNs), Transfer Learning with Pretrained Networks, Ensemble Models

Feature Engineering & Transformation

Resizing	+	Normalization	+	Data Augmentation	+	Encoding Labels
256(W) X 256(D) X 3(RGB)		Adjusted pixel values to 0-1		Rotation Rage: 20, Width Shift Range: 0.2, Height Shift Range: 0.2 Shear Range: 0.2, Zoom Range: 0.2 Horizontal Flip: True, Fill Mode: Nearest		Ballet Flat (1,0,0,0,0) Boat (0,1,0,0,0) Brogue (0,0,0,1,0) Clog (0,0,0,1,0) Sneaker (0,0,0,0,1)

1. Resizing standardizes image dimensions to ensure uniformity of input for the model.
2. Normalization adjusts pixel values from 0-255 to 0-1 to prevent disruption or slowdown in the learning process due to large integers.
3. Data Augmentation enhances dataset diversity through random, realistic image transformations, facilitating better model generalization.
4. Encoding Labels converts categorical features into numerical ones, a necessary step for many machine learning algorithms.

Model Development

Baseline Model: CNN with Max Pooling

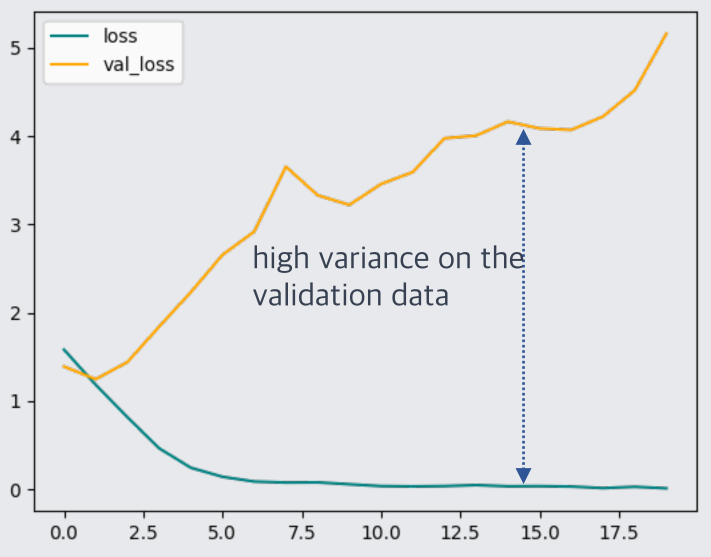
- This project has adopted Nicholas Renotte's binary image classifier, as presented in his 'Build a Deep CNN Image Classifier with ANY Images' tutorial. His baseline model was expanded for multi-class image classification, incorporating enhancements like data augmentation and transfer learning for improved performance.

Layer (type)	Output Shape	Param #	Hyper Parameters
conv2d (Conv2D)	(None, 254, 254, 16)	448	Filters: 16, Filter size (3X3), Stride: 1, Activation: 'relu'
max_pooling2d (MaxPooling2D (None, 127, 127, 16) 0)	(None, 127, 127, 16)	0	
conv2d_1 (Conv2D)	(None, 125, 125, 32)	4640	Filters: 32, Filter size (3X3), Stride: 1, Activation: 'relu'
max_pooling2d_1 (MaxPooling (None, 62, 62, 32) 0 2D)	(None, 62, 62, 32)	0	
conv2d_2 (Conv2D)	(None, 60, 60, 64)	18,496	Filters: 64, Filter size (3X3), Stride: 1, Activation: 'relu'
max_pooling2d_2 (MaxPooling (None, 30, 30, 64) 0 2D)	(None, 30, 30, 64)	0	
flatten (Flatten)	(None, 57,600)	0	
dense (Dense)	(None, 256)	14,745,856	Nodes: 256, Activation: 'relu'
dense_1 (Dense)	(None, 5)	1,285	Activation: 'softmax'

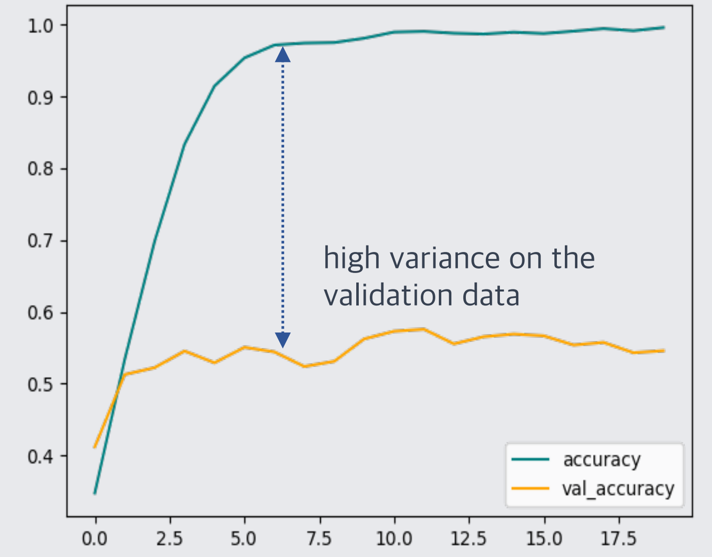
Total params: 14,770,725, Trainable params: 14,770,725, Non-trainable params: 0

Model Development

Validation Loss



Validation Accuracy



Epochs 20, Validation Set 25%

Baseline Model Summary	
Precision	0.54825294
Recall	0.54238683
Accuracy	$\boxed{0.54485595}$

Low accuracy: high variance on the test data

class	precision	recall	f1-score	support
Ballet Flat	0.38	0.41	0.39	97
Boat	0.45	0.34	0.39	236
Brogue	0.50	0.95	0.66	192
Clog	0.64	0.49	0.56	424
Sneaker	0.62	0.56	0.59	266

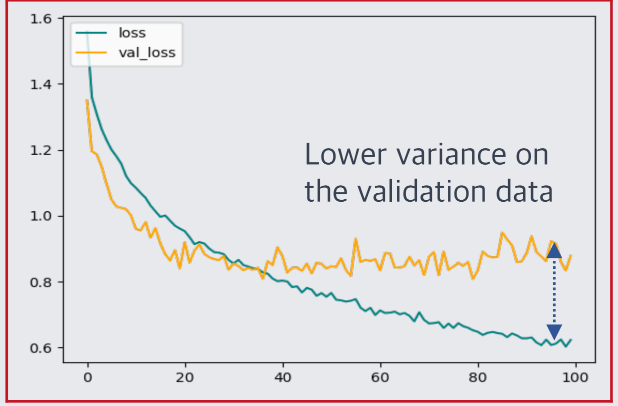
Our current model exhibits a significant overfitting issue, as evidenced by its unsatisfactory accuracy.

Possible Solutions: data augmentation, regularization (early stopping), and transfer learning strategies

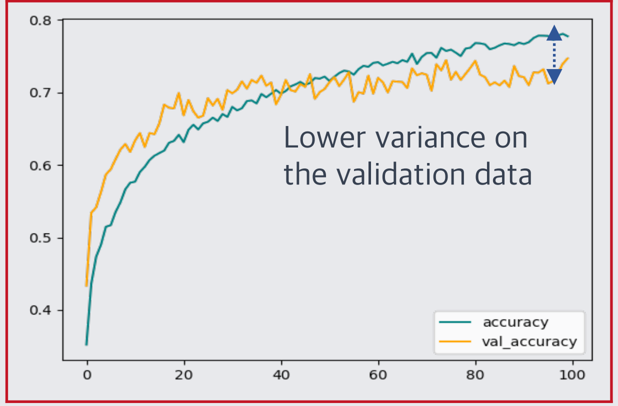
Optimal Model Selection: Data Augmentation

Data Augmentation Only

Validation Loss

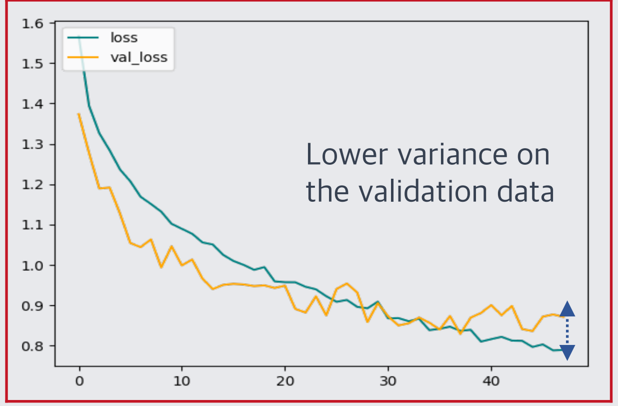


Validation Accuracy

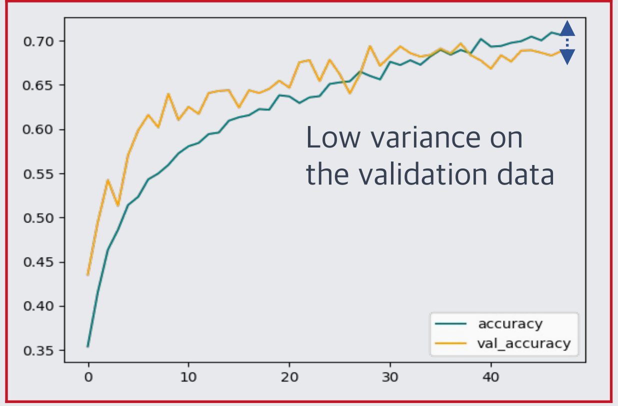


Data Augmentation + Early Stopping

Validation Loss



Validation Accuracy



Epochs 100, Validation Set 25%, Early Stopping Patience 10

By augmenting the training data with diverse transformations and leveraging pre-trained features, this strategy enhances the model's ability to generalize and mitigate overfitting more effectively than using each technique separately.

Model 2 Summary ***Data Augmentation + Early Stopping***	
Precision	0.6798246
Recall	0.51028806
Accuracy	0.6090535

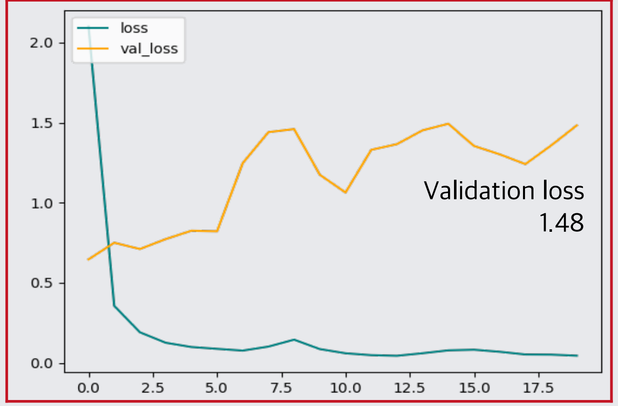
While addressing the overfitting issue, the accuracy has shown marginal improvement necessitating further investigation for potential enhancements.

class	precision	recall	f1-score	support
Ballet Flat	0.49	0.62	0.55	97
Boat	0.47	0.56	0.55	236
Brogue	0.58	0.73	0.65	192
Clog	0.76	0.54	0.64	424
Sneaker	0.66	0.66	0.66	266

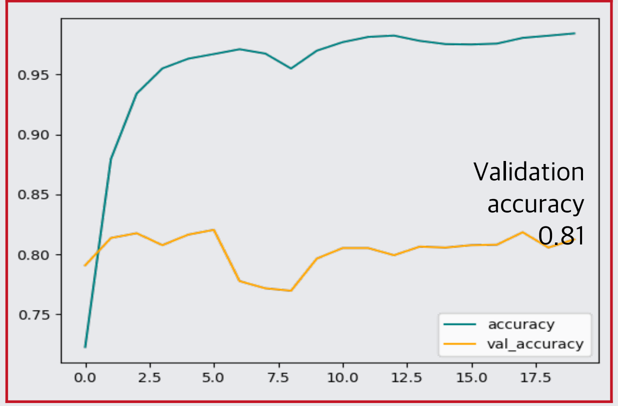
Optimal Model Selection: Transfer Learning

ResNet152V2

Validation Loss

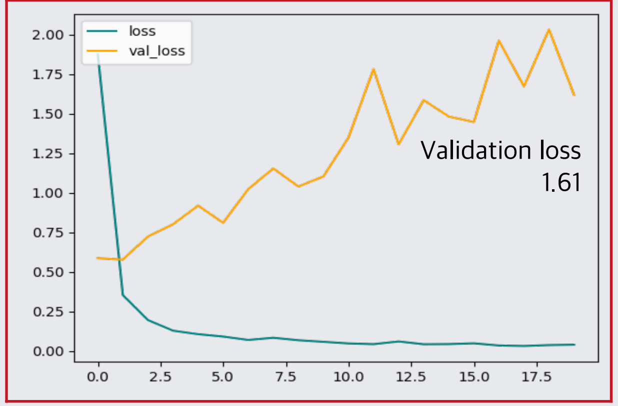


Validation Accuracy

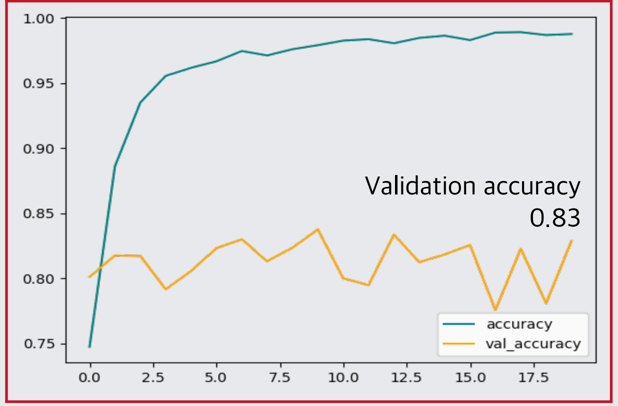


DenseNet201

Validation Loss



Validation Accuracy



Epochs 20, Validation Set 25%

DenseNet201 was selected for its slightly better performance in precision, recall, and accuracy. There was a significant accuracy improvement of +29% compared to the baseline model. However, there has been a subsequent increase in variance.

Model 3 Summary ***DenseNet201***	
Precision	0.83568466
Recall	0.8288066
Accuracy	0.8312757

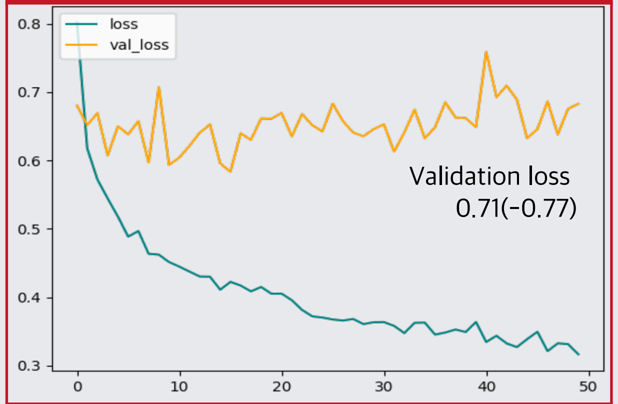
Transfer learning with DenseNet201 significantly improved accuracy, but overfitting needs to be addressed.

class	precision	recall	f1-score	support
Ballet Flat	0.77	0.76	0.77	97
Boat	0.74	0.76	0.75	236
Brogue	0.76	0.98	0.86	192
Clog	0.91	0.83	0.87	424
Sneaker	0.88	0.91	0.84	266

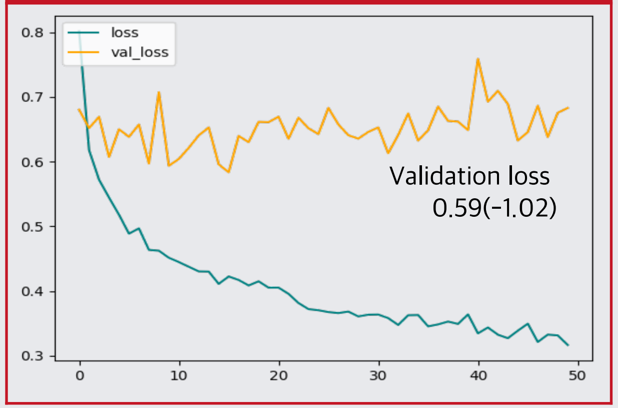
Optimal Model Selection: Combined Strategy

ResNet152V2 + Data Augmentation	DenseNet201 + Data Augmentation

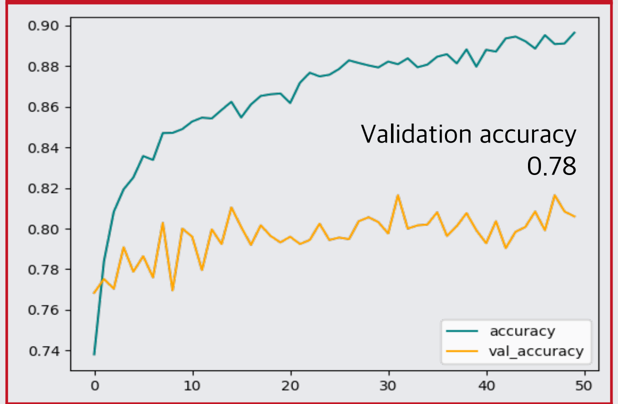
Validation Loss (ResNet152V2 + Data Augmentation)



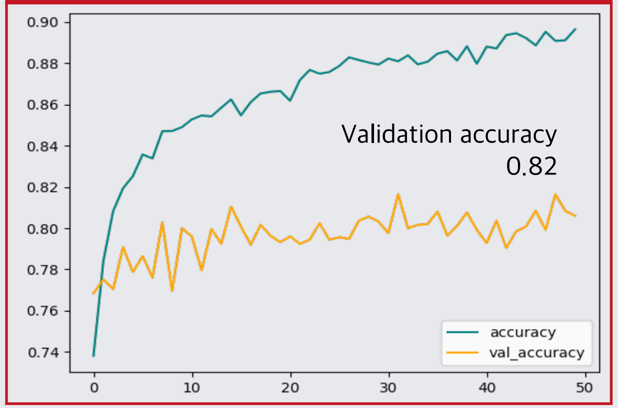
Validation Loss (DenseNet201 + Data Augmentation)



Validation Accuracy (ResNet152V2 + Data Augmentation)



Validation Accuracy (DenseNet201 + Data Augmentation)



Epochs 50, Validation Set 25

DenseNet201 achieved better results compared to ResNet152, exhibiting higher precision, recall, and accuracy. It also demonstrated improved generalization capabilities, reducing validation loss from 1.61 to 0.59 when trained with augmented data.

Model 4 Summary ***DnseNet201 + Data Augmentation***	
Precision	0.84273505
Recall	0.81152266
Accuracy	0.82222223

Although the model demonstrated improved generalization, there was a slight drop in accuracy by 0.009. This suggests the need to explore alternative modeling methodologies.

class	precision	recall	f1-score	support
Ballet Flat	0.71	0.75	0.73	97
Boat	0.77	0.67	0.72	236
Brogue	0.74	0.97	0.84	192
Clog	0.91	0.84	0.87	424
Sneaker	0.86	0.85	0.86	266

Optimal Model Selection: Ensemble Model

Ensemble Model Prediction Results for Unseen Data

Actual	Ballet Flat	Boat	Brogue	Clog	Sneaker
Prediction	Ballet Flat	Boat	Brogue	Clog	Sneaker

Actual	Ballet Flat	Boat	Brogue	Clog	Sneaker
Prediction	Ballet Flat	Boat	Brogue	Clog	Sneaker

Actual	Ballet Flat	Boat	Brogue	Clog	Sneaker
Prediction	Ballet Flat	Boat	Brogue	Clog	Sneaker

The unseen data is not included in the training, test, or validation sets.

The ensemble model*, combining DenseNet201 and ResNet152V2 trained with augmented data, successfully addressed the overfitting issue. The accuracy improved by 0.027, reaching 0.85, and when evaluated on unseen data, the model achieved a perfect prediction accuracy of 100%, demonstrating strong generalization capabilities.

Model 5 Summary ***Ensemble Model****	
Precision	0.8562660331279
Recall	0.8493827160493
Accuracy	0.8493827160493

By applying transfer learning to a model trained on augmented data, we were able to enhance accuracy through the implementation of ensemble modeling.

class	precision	recall	f1-score	support
Ballet Flat	0.73	0.82	0.77	97
Boat	0.79	0.77	0.78	236
Brogue	0.78	0.98	0.87	192
Clog	0.93	0.84	0.89	424
Sneaker	0.89	0.85	0.87	266

THE UNIVERSITY OF CHICAGO

Performance Analysis: Final Model

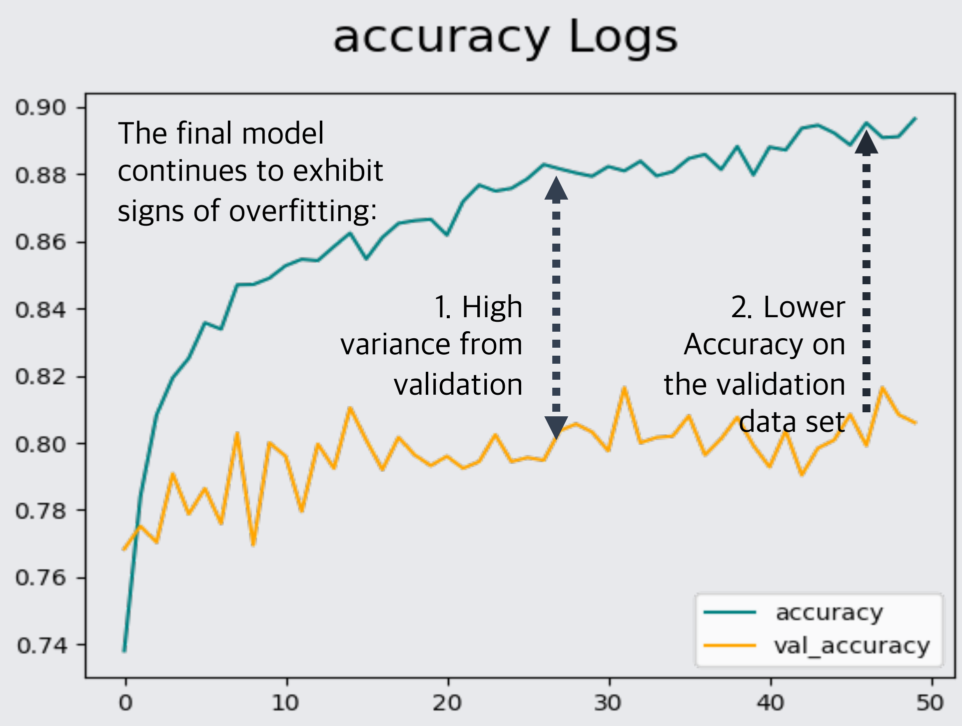
Final Model Summary	
Topic	Description
Methodologies	Ensemble Learning, Transfer Learning, Data Augmentation
Data Augmentation Techniques	Rescaling, Rotation (±20 degrees), Width Shift (20% of width), Height Shift (20% of height), Shearing (20% intensity), Zooming (20% intensity), Horizontal Flipping
Base Models	DenseNet201, Resnet152V2 (Transfer Learning)
Data Splitting	Training set: 10,000, Validation set: 2,500 (25%)
Training Parameters	Epochs: 50
Loss Function	Categorical Cross-Entropy

From the baseline CNN model, the ensemble model enhanced precision, recall, and accuracy each by a significant 0.3. Overfitting, a primary concern, was strategically mitigated through the application of data augmentation techniques.

Final Model Summary Ensemble Model	
Precision	0.8562660331279
Recall	0.8493827160493
Accuracy	0.8493827160493

class	precision	recall	f1-score	support
Ballet Flat	0.73	0.82	0.77	97
Boat	0.79	0.77	0.78	236
Brogue	0.78	0.98	0.87	192
Clog	0.93	0.84	0.89	424
Sneaker	0.89	0.85	0.87	266

Performance Analysis



Given the content quality issues present in the data, the final model is currently affected by substantial variance and bias. Therefore, further refinements should be implemented to enhance model performance.

Insufficient Categories

The current categorization of certain items, such as classifying some women's sandals as ballet flats, may be too broad. This can lead the model to learn overly generalized representations of features within a category, contributing to higher variance in the predictions.

Miscategorization

Incorrect labeling of items, such as miscategorizing sneakers as boat shoes, can distort the model's understanding of features associated with each category. This can cause the model to develop incorrect patterns, which may increase the bias in predictions.

Presence of Irrelevant Images

The dataset contains some completely unrelated images within certain categories. This can cause noise in the data, and the model might overfit to these irrelevant features, thereby increasing the variance and reducing its ability to generalize on unseen, relevant data.

Performance Analysis

Strategies for Improving Accuracy in the current model

Data Quality and Content

Remove irrelevant images. Misleading images can confuse the model and reduce its ability to correctly identify patterns, decreasing accuracy.

Increase the number of classes because the current categories are too broad. More specific classes can help the model differentiate between similar items, thereby improving its accuracy.

Understanding Transfer Learning Models

Favor DenseNet201, given its performance. Its architecture supports better gradient flow and feature reuse, potentially improving the accuracy on the shoe image dataset.

EfficientNetB1 showed limited performance, perhaps due to the varied image quality. It's designed to perform optimally on larger datasets and with longer training periods.

Model Fine-Tuning

Leverage a variety of data augmentation techniques. Apart from the basic transformations, use more advanced methods such as brightness and contrast adjustment, random cropping, or even synthetic data generation.

Implement learning rate scheduling and hyperparameter tuning to refine the model's learning process and achieve optimal accuracy.

Performance Analysis: Final Model

Initial CNN Model

- Exhibited high variance, indicative of overfitting.
- Demonstrated poor generalization on unseen data.

Application of Data Augmentation

- The model showed improved generalization, able to handle unseen data better.
- However, the overall accuracy was still unsatisfactory.

Application of Transfer Learning

- The model's accuracy was significantly enhanced.
- However, this also led to an increase in variance.

Transfer Learning + Data Augmentation

- This combination effectively decreased the variance, providing a more balanced model.

Final Ensemble Model

- Further increased the overall accuracy of predictions
- To boost model accuracy, consider refining data quality, optimizing transfer learning models, expanding data augmentation, and enhancing fine-tuning strategies.

Future Work

Model Expansion

Plan to increase the model's scope to identify additional shoe types for commercial application, with the ultimate aim of extending it to all fashion items. This expansion necessitates a more complex model architecture.

Regularization Strategies

As fashion designs, including shoes, often evolve rapidly with trends, regularization is key to maintaining model relevance. Techniques such as L1/L2 regularization, Dropout, and Early Stopping will be employed to improve the model's generalization capabilities.

Data Acquisition and Augmentation

In order to adapt to the fast-paced nature of fashion trends and enhance the model's robustness, there will be a concerted effort to increase data collection and apply data augmentation techniques for improved model performance.

References

1. Shoe Classification Dataset: https://www.kaggle.com/datasets/utkarshsaxenadn/shoes-classification-dataset-13k-images

2. Binary Image Classifier Tutorial: https://www.youtube.com/watch?v=jztwpsIzEGc

3. Problem Statement: https://sell.amazon.com/start?ld=SEUSSOAGOOG-NAG047B-EGC-D

4. Data Assumption
   - https://www.fiverr.com/joyvijayp/make-tech-pack-for-your-footwear-design
   - https://www.oxfordlearnersdictionaries.com/us/definition/english/shoe_1
   - https://www.nike.com/t/dunk-high-retro-mens-shoes-dTVTCk/DD1399-105?nikemt=true&srsltid=AR57-fBWR2UtQE0HCLCZHgaXhGxNmPxDJzO_sO7_8Tz6pNiBdr6qWwcfBq4

5. Images for Model Estimation (continued)
   - Ballet Flat : https://www.repetto.com/us/camille-ballerinas-flammy-red-v511v-550-us.html#&gid=null&pid=3
   - Ballet Flat : https://margauxny.com/collections/the-demi?color=Black&variant=37228992069801
   - Ballet Flat: https://www.coach.com/products/whitley-mary-jane/CG603-BLK
   - Ballet Flat: https://www.alexandradecurtis.com/shoes/livia-ballet-flats-blush-pink
   - Ballet Flat : https://www.everlane.com/products/womens-italian-leather-day-ballet-flat-andorra
   - Boat: https://flawlesscrowns.wordpress.com/2011/05/11/fred-perry-dury-canvas-boat-shoe/
   - Boat: https://www.sperry.com/en/starfish-1-eye-palm-embossed-boat-shoe/56478W.html
   - Boat: https://www.rancourtandcompany.com/products/read-boat-shoe-natural-chromexcel
   - Boat: https://www.timberland.com/shop/womens-noreen-3-eye-lug-handsewn-shoe-brown-tan-51304214

References

4. Images for Model Estimation

- Boat: https://thehelmclothing.com/products/leather-boat-shoes
- Boat: https://us.dockers.com/collections/mens-shoes/products/vargas-boat-shoes-s90304091
- Brogue: https://bananarepublicfactory.gapfactory.com/browse/product.do?pid=422844
- Brogue: https://www.beckettsimonon.com/products/yates-oxford?variant=29480266268751&
- Brogue: https://www.drmartens.com/us/en/3989-yellow-stitch-smooth-leather-brogue-shoes/p/22210001
- Brogue: https://www.hawesandcurtis.com/mens-tan-leather-derby-brogue-shoe-oepjx152-z03
- Brogue: https://www.church-footwear.com/us/en/men/collection/brogues/products.Polished_Binder_Oxford_Brogue.EEB002_9XV_F0AAB_F_000000.html
- Clog: https://www.birkenstock.com/us/boston-suede-leather/boston-suede-suedeleather-softfootbed-eva-u_46.html
- Clog: https://www.bottegaveneta.com/en-us/beebee-clog-black-812884005
- Clog: https://www.amazon.com/crocs-Unisex-Classic-Black-Women/dp/B0014BYH62
- Clog: https://www.jcrew.com/us/p/womens/categories/shoes/clogs/convertible-leather-clogs/
- Sneaker: https://www.adidas.com/us/samba-og-shoes/B75806.html?
- Sneaker: https://www.finishline.com/store/browse/productDetail.jsp?productId=prod2823333&brand_name=NIKE&styleId=DD1869&colorId=103
- Sneaker: https://www.tower-london.com/en-us/products/new-balance-530-womens-white-natural-indigo-trainers
- Sneaker: https://www.amazon.com/adidas-Originals-Womens-Forum-Sneaker/dp/B08MK43YMJ