# Towards Fashion Image Annotation: A Clothing Category Recognition Procedure

Tryfon-Rigas Tzikas
tzikasta@ece.auth.gr
Electrical and Computer Engineering,
Aristotle University of Thessaloniki
Thessaloniki, Greece

Alexandros-Charalampos Kyprianidis
alexandros.kyprianidis@issel.ee.auth.gr
Electrical and Computer Engineering,
Aristotle University of Thessaloniki
Thessaloniki, Greece

Maria Kotouza
maria.kotouza@issel.ee.auth.gr
Electrical and Computer Engineering,
Aristotle University of Thessaloniki
Thessaloniki, Greece

Sotirios-Filippos Tsarouchis
sotiris.tsarouchis@issel.ee.auth.gr
Electrical and Computer Engineering,
Aristotle University of Thessaloniki
Thessaloniki, Greece

Antonios Chrysopoulos
achryso@issel.ee.auth.gr
Electrical and Computer Engineering,
Aristotle University of Thessaloniki
Thessaloniki, Greece

Pericles Mitkas
mitkas@auth.gr
Electrical and Computer Engineering,
Aristotle University of Thessaloniki
Thessaloniki, Greece

## Abstract

In contemporary clothing industry, design, development and procurement teams are constantly asked to present more products with fewer resources in a shorter time. Thus, clothing companies that aim to remain competitive in today's market have to deploy new Artificial Intelligence techniques aiming at the automation of their traditional procedures. In this direction, the presented approach utilizes a deep learning model to accurately classify fashion images. The predictions are intended to be used on a personalized recommendation system, that acts as an assistant for the fashion designers. Two well established architectures are studied, VGG and ResNet, as well as a variation of ResNet. The realized experiments include: (a) architecture comparison, (b) hyperparameter tuning and classification, and (c) transfer learning. Two fashion datasets are used for the model training and classification: DeepFashion (for training the model from scratch) and iMaterialist (used to evaluate the transferability of the produced model). The results show that the first set of experiments achieved 80.5% accuracy, whereas the pre-trained model used on the second dataset led to a decrease of 60% on training time, while attaining satisfying results.

***CCS Concepts:*** • **Computing methodologies** → **Object recognition**; *Supervised learning by classification*; Neural networks; • **Applied computing** → Consumer products.

***Keywords:*** object classification, fashion clothing images, fine-tuning, convolutional neural networks

## 1 Introduction

Fashion clothing is one of the oldest industries, occupying one of the highest market shares. In this age of fast fashion, trends change in a highly frequent manner, making it an appropriate field for applying optimization techniques to efficiently extract valuable information from the huge amount of generated data. To this end, contemporary clothing brands tend to introduce Artificial Intelligence (AI) techniques, aiming to improve the processes of supply chain, while keeping up to date with the newest fashion trends. Fashion houses such as Hugo Boss[1] and Tommy Hilfiger[2] have already developed AI-driven tools to improve the design process, whereas Prada[3] uses AI to deliver high-quality content faster.

The development of such tools was not feasible before the evolution of Deep Learning and Computer Vision: image recognition, detection, segmentation and generation, as well as 3D reconstruction, are some of the techniques that are being used in the development of fashion related solutions. The emergence of an abundance of related projects is justified by the rapid growth in the specific scientific fields.

In this paper, Deep Learning algorithms for clothing category classification are evaluated. Two datasets are used as inputs, DeepFashion and iMaterialist, while data augmentation techniques are applied on them. The first one is used to train the model from scratch, while the second one to evaluate the transferability of the produced model. The models that were used during the experiments are VGG16, ResNet50 and a variation of ResNet50.

*AI4FASHION2020, September 02–04, 2020, Athens, Greece*


[1] https://www.hugoboss.com/fashionstories/digitalisation-is-and-remains-a-big-trend-which-has-already-been-embraced-by-hugo-boss/fs-story-1e6xd6hk2kr8e.html
[2] https://www.ibm.com/blogs/think/2018/01/tommyhilfiger-ai/
[3] https://www.pradagroup.com/en/news-media/news-section/prada-group-expands-collaboration-with-adobe.html

The proposed solution is part of the Data Annotation module introduced in our previous work [11], where an AI-enabled system utilized towards the improvement of clothing design process was proposed. Specifically, the aforementioned system is responsible for retrieving, organizing and combining data from many different sources, while taking into account the designers' preferences, in order to suggest clothing products of interest and help fashion designers with the decision-making process.

The rest of paper is organized as follows. Section 2 lists related works. Section 3 introduces the methodology. Section 4 presents the experimental setup, datasets and results. Section 5 contains the conclusion and future work.

## 2 Related Work

Several research works have been realized in the field of AI-enabled Fashion applications. There are many works that tried to discern the AI applications in the fashion industry in four categories [7]: (a) apparel design, (b) manufacturing, (c) retailing, (d) supply chain management. In the work of [13] a comprehensive review of AI systems in apparel supply chains is presented, while in [5] an empirical review on existing apparel recommendation systems is conducted.

Fashion image analysis has emerged as a challenging task. The majority of the approaches that have been used over time can be described as follows: (a) traditional features learning methods based on manually created features which are then processed by machine learning algorithms [15], (b) Deep Learning algorithms based on deep neural networks and especially convolutional neural networks. In most cases, the models that have been developed achieve high results concerning image classification and recognition. [12] [3] [9]

In the area of fashion image classification, Hidayati et al. [9] proposed a classification technique that recognizes clothing genres based on visually differentiable style elements. Additionally, Cychnerski et al.[2] presented a set of experiments in order to evaluate ResNet and SqueezeNet.

Many datasets have been introduced as test-beds to apply various AI techniques in the field of fashion. DeepFashion [12] is composed of 800,000 images which are richly annotated with attributes, clothing landmarks and correspondence of images taken under different scenarios. DeepFashion2 [3] is an improved version of DeepFashion, with enriched annotations; style, scale, viewpoint, occlusion, bounding box and dense landmarks were added.

## 3 Methodology

The clothing category classification, as well as the fine-tuning of an existing model to another dataset are challenging tasks. In Figure 1, the proposed approach is described, being divided in three steps. As a first step, three different deep learning architectures are tested: (a) VGG16, (b) ResNet50 and (c) a variation of ResNet50 (ResNet50v2), by using the DeepFashion dataset, after applying image pre-processing techniques. The next step contains the selection of the architecture with the highest accuracy, by performing a grid search for the image augmentation parameters and the model's training hyperparameters. In the last step, the fine-tuned model is used on the iMaterialist dataset, to evaluate the transferability of the produced model.

### 3.1 Image Pre-processing

The efficiency of the model is heavily dependent on the input dataset that is used during the training process. Taking this into consideration, the images need to be cropped, using the provided bounding boxes from the dataset, to exclude non-relatable objects as well as background noise, in order to restrain the model from capturing irrelevant information. Moreover, in a multi-class classification problem, each image corresponds to one label, thus it needed to avoid having multiple clothes in a single image, as it can mislead the training process and affect its performance in a negative manner.

In order to achieve higher performance and reduce overfitting, Data Augmentation techniques are applied, on the available training set, in the following order: 1) rotation, 2) shearing, 3) horizontal flip and 4) zoom in or out; experimenting on each one of them to fine-tune them. Starting with the first technique, a range of low values was tested and the optimal values were kept in the end.

### 3.2 Clothes Recognition with ResNet

There are many state-of-the art solutions in the literature related to image recognition using Deep Learning techniques. Architectures like VGG [14] and ResNet [10] are proved to be ideal for recognizing clothing categories from fashion images [1] [2]. More specifically, VGG16 and ResNet50 are commonly used in this field.

In this work, experimentation with VGG16 and ResNet50 was realized. Additionally, a variation of ResNet50 was investigated, which is characterized by an architecture with the following modifications in the skip connection: the batch normalization and the ReLU function takes place before the convolutional layer [2]. This variation of ResNet50 was chosen as the one with the best performance amongst other variation attempts on the input dataset.

### 3.3 Hyperparameter Tuning

Hyperparameter tuning is a crucial task towards achieving the optimal performance in Deep Learning modelling. In this process, a set of optimizers were investigated in order to find the appropriate one for the problem at hand. More specifically, the optimizers examined are Adam, Adadelta, Adamax, Adagrad, SGD.

Weight initialization of a Deep Learning network strongly affects the performance of the model, since problems like vanishing and exploding gradients are tackled by using the

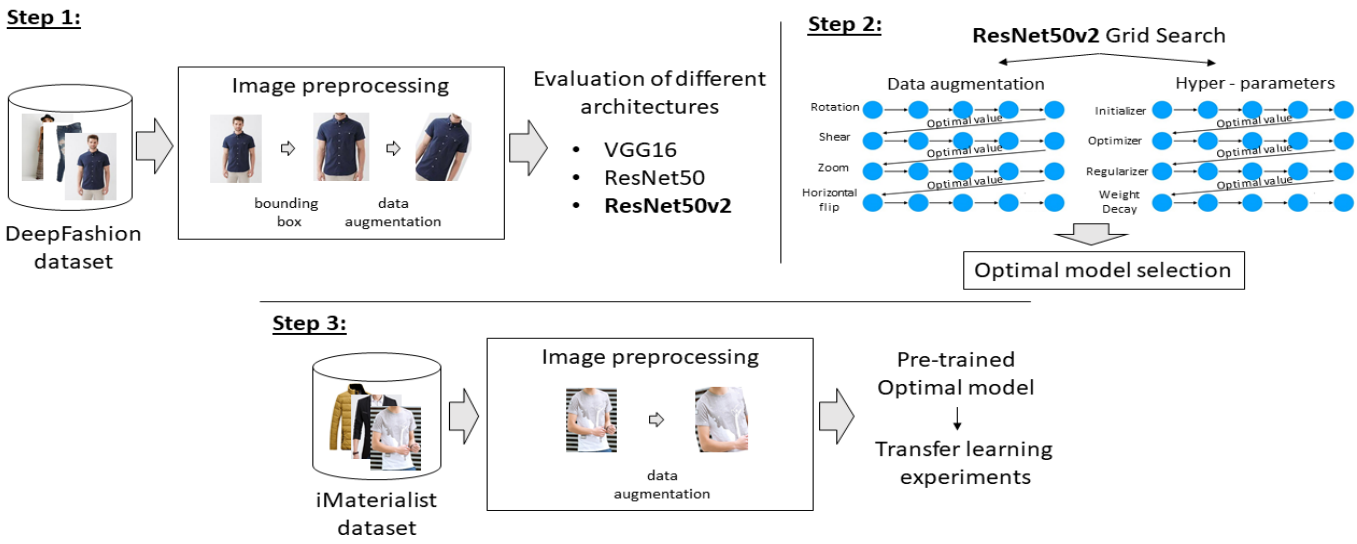


**Figure 1.** Overview of the proposed methodology

correct initializer. The following initializers were used in the experiments: (a) Random Normal, (b) He Normal [8], (c) Glorot Normal [4], (d) Zeros, (e) He Uniform [8], and (f) Glorot Uniform [4].

In addition, regularization restricts the exponential growth of model's weights and prevents the model from overfitting. The techniques employed in the proposed approach are a combinations of regularizers and weight decay. Both these parameters are investigated in regard with the learning rate, as they are correlated with it. The regularizers examined are as follows: (a) L1 (b) L2 (c) L1 & L2, while the weight decay values are: (a) 0.98, (b) 0.95, (c) 0.75.

### 3.4 Transfer Learning

After the completion of the first set of experiments, focused on the multi-class classification problem of clothing categories, we proceed with the examination of the second set, which deals with the evaluation of the performance of an already trained model in another dataset, making use of transfer learning techniques. The evaluation of the model in a second dataset can be broken down in two cases: (a) evaluating the pre-trained model without further training, and (b) using the pre-trained model as a starting point to re-train either the whole model, or only specific layers. The whole idea is based on the similarity between the two fashion datasets and on the fact that they share common low-level features, which are also captured from the weights of the bottom layers of the model. The main hypothesis should improve the model's performance as it can achieve comparative results in significant less time.

## 4 Experiments

This section contains the experimental process on the problem of multi-class clothing categories classification and the evaluation of the produced models' performance. The section is composed of three sets of experiments, as follows: (1) architecture comparison, (2) hyperparameter tuning and classification, and (3) transfer learning.

### 4.1 Datasets

Two datasets were used for the training and evaluation of the models, DeepFashion and iMaterialist. DeepFashion dataset [12] consists of 800,000 images characterized by many features and labels. iMaterialist dataset [6] consists of 1,000,000 images and contains 8 groups of 228 fine-grained attributes. The imbalanced distribution of the classes in each dataset was balanced by randomly choosing 5000 images for every clothing category, using 50.000 images in total. They were split into training, validation and test set with ratios of 0.7, 0.15, 0.15, respectively.

### 4.2 Experimental Setup

Input images were scaled down to 224x224 RGB images and classified into 10 classes including *coat and jacket*, *dress*, *top*, *shorts*, *trousers*, *skirt*, *leggings and jeggings*, *outfit*, *special occasion* and *suits*. The models were trained on a Nvidia Tesla K40c GPU with 32GB memory RAM and utilizing an Intel Xeon E5-2630 processor. The batch size that was used during training is 32 and the initial learning rate was set according to Keras defaults values for each optimizer (0.01 for SGD and 0.001 for the rest of them).

### 4.3 Results

**4.3.1 Architecture Comparison.** The architectures tested for the classification of the provided clothing categories are the following: VGG16, ResNet50 and a variation of ResNet50 (ResNet50v2) [2]. They were all tested using the same values on each hyperparameter, based on the configuration in Table

1. Moreover, Table 1 makes clear that ResNet50v2 outperforms the rest of the models, achieving accuracy 74%; thus it is selected to be used for the rest of the experiments.

The performance of the models was measured with the usage of the following evaluation metrics: accuracy, precision, recall and f1 score.

**Table 1.** Model initialization parameters and architecture comparison

| Parameters | Values | Model | Accuracy |
|---|---|---|---|
| Optimizer | Adam | VGG16 | 67% |
| Initializer | Glorot Uniform | | |
| Learning Rate | 0.01 | ResNet50 | 70% |
| Weight Decay | 0.9 | | |
| Regularizer | L1 | ResNet50v2 | 74% |
| Image Augmentation | None | | |

**4.3.2 Classification Results.** Towards the improvement of the produced model's performance, many experiments were conducted in order to find the best configuration of the available hyperparameters. During this process, a grid search for the image augmentation parameters was performed, as well as the model's training hyperparameters, in order to boost the accuracy of the model. The order in which the experiments were performed is as follows: (a) Image augmentation (b) Initializer, (c) Optimizer, (d) Learning rate and Regularizer, (e) Learning rate and weight decay. In the following experiments the default parameters are used for the initial configuration, as mentioned in Table 1. The order in which each parameter's experiments are conducted is important, as with the completion of each one, the optimal value of the corresponding parameter is extracted and is used in the configuration of the following experiments.

The results of the image augmentation experiments, are presented in Table 2. The optimal values for each technique are the following: (a) *Rotation: 10*, (b) *Shear: 0.2*, (c) *Zoom: 0.05* and (d) *Horizontal flip: True*. The optimal values led the produced model to not only achieve better performance, but to avoid overfitting, as well. It is clear that the model performs better when the image augmentation process causes mediocre changes in the datasets.

In Table 3, the results of the various initializers and optimizers are presented. In the first case *Glorot Normal* achieved the best results, while *Zeros* provided the worst, as expected. As far as the optimizers are concerned, they all achieved similar results, except from *SGD*. The reason behind this is that *SGD* demands additional fine-tuning to determine the appropriate hyperparameters, in contrast with the rest of the optimizers, who are adaptive gradient methods. Among the optimizers, *Adadelta* achieved the highest accuracy.

The results of the experiments conducted in order to determine the weight decay and regularizer are presented in Table 4. Their optimal values are strongly dependent on the learning rate parameter. For this reason each parameter is tested in respect to different values of learning rate. The best values of three parameters coming in pairs are as follows: (a) *learning rate: 1, regularizer: L2*, (b) *learning rate: 1, weight decay: 0.75*.

**Table 2.** Image augmentation experiments

| Rotation | Accuracy | Shear | Accuracy |
|---|---|---|---|
| 0 | 71.0% | 0 | 73.0% |
| 10 | 73.0% | 0.05 | 76.0% |
| 30 | 67.0% | 0.1 | 71.0% |
| 90 | 52.0% | 0.2 | 77.0% |
| **Zoom** | **Accuracy** | **Horizontal Flip** | **Accuracy** |
| 0 | 77.0% | True | 77.5% |
| 0.05 | 77.2% | | |
| 0.1 | 76.0% | False | 77.2% |
| 0.2 | 74.0% | | |

**Table 3.** Initializer and optimizer experiments

| Initializer | Accuracy | Optimizer | Accuracy |
|---|---|---|---|
| Random Normal | 78.3% | Adam | 78.0% |
| He Normal | 77.8% | Adagrad | 77.8% |
| Glorot Normal | 78.8% | Adadelta | 80.0% |
| Zeros | 10.0% | SGD | 71.0% |
| He Uniform | 77.6% | Adamax | 79.0% |
| Glorot Uniform | 77.5% | | |

**Table 4.** Learning rate, weight decay and regularizer experiments

| | Regularizer | | | Weight Decay | | |
|---|---|---|---|---|---|---|
| **Learning rate** | L1 | L2 | L1 & L2 | 0.98 | 0.95 | 0.75 |
| 0.01 | 67% | 63% | 68% | 67% | 63% | 60% |
| 0.1 | 65% | 78% | 78% | 76% | 78% | 76% |
| 1 | 73% | 80% | 75% | 80% | 80% | 81% |

**Table 5.** Model optimization parameters

| Image Augmentation | Values | Parameters | Values |
|---|---|---|---|
| Rotation | 10 | Optimizer | Adadelta |
| Shear | 0.2 | Initializer | Glorot Normal |
| Zoom | 0.05 | Learning Rate | 1 |
| Horizontal Flip | True | Weight Decay | 0.75 |
| | | Regularizer | L2 |

The final trained model using the optimal parameters achieved 80.5% accuracy, as presented in Table 5. Figure 2 is the confusion matrix of the model for each class. The diagonal of the matrix presents the true positive value per class. The classes *Skirt*, *Trousers*, *Dress* and *Shorts* are classified better than the rest, while many samples of *Outfit* and *Suits* are misclassified as *Coat* and *Dress* respectively, since there is a vivid resemblance between the images of these classes.

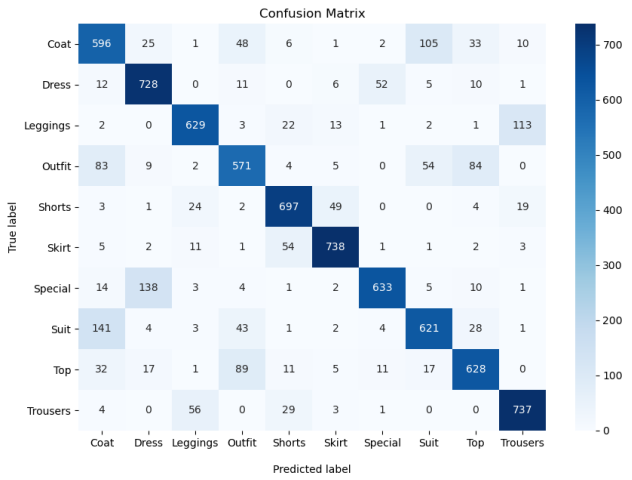


**Figure 2.** ResNet50v2 evaluated in DeepFashion

**4.3.3 Transfer Learning Results.** In this section, the performance of the Deep Learning model produced from the first set of experiments is evaluated on the iMaterialist dataset, which was not used previously. The datasets have many visual features in common, as they both are used for classifying fashion clothing images to categories. Therefore, it is assumed that the pre-trained model can be used as a baseline, upon which we can apply a set of slight weight adjustments through fine-tuning to improve its performance, while using a low value for the training learning rate. In order to have comparative results, the same hyperparameters and the evaluation results of the pre-trained model in iMaterialist were maintained as benchmark in the fine-tuning experiments.

Table 6 contains the comparison results of the fine-tuning experiments against the ones achieved by the pre-trained model, which is the benchmark and has not undergone any further training. The differentiation between the experiments lies on the model's layers that each time are trained. Thus, for the first step of the Transfer Learning process the pre-trained model was applied on the input dataset as is, without changing any of the pre-defined hyperparameters. The results were very poor, since the model achieved a mere 38% accuracy, indicating that the two datasets contain different content and they cannot be processed by the produced model without additional training.

**Table 6.** Transfer Learning Experiments

| Experiments | Precision | Recall | F1 Score | Accuracy |
|---|---|---|---|---|
| Benchmark (No training) | 40.7% | 37.8% | 37.3% | 38.0% |
| Last layer | 46.3% | 42.4% | 42.1% | 42.0% |
| Whole model | 65.2% | 64.6% | 64.7% | 62.5% |
| Without pre-trained weights | 65.1% | 64.9% | 64.8% | 65.0% |

On the second step of the experimental process, all the layers of the model were frozen, except from the last one, in order to keep the learned features intact and modify only the classifier's weights, which constitutes the last layer of the model. The results show a slight improvement over the benchmark on each evaluation metric.

To further improve the model's performance on the new dataset, the whole model was unfrozen, which actually led to significantly better results. The model achieved 62.5% accuracy, almost 20% better than the previous best performance, revealing that even though the datasets share common features, as they both contain fashion clothing images, they also appear to have variant inputs.

To highlight this last point, the confusion matrix of the last experiment is presented on Figure 3. The classes *Shorts*, *Trousers*, *Coat* are classified with greater confidence, while *Leggings* are misclassified as *Trousers* and *Dress* as *Skirts* and vice versa. This behavior may derive from either annotation fault or the fact that these two classes share many visual characteristics, as a long skirt can be easily misjudged as a dress.

Lastly, the model was trained from scratch, without using any weights originating from the pre-trained model. The model achieved 65% accuracy, surpassing the previous results. The result is completely justified, as the newly estimated hyperparameters are more suitable for whole model training, while in fine-tuning it is needed to use lower learning rate to slightly adjust the weights. Comparing the performance of the model trained from scratch and the model trained using the pre-trained weights, it seems that the second one achieved 2.5% less accuracy. However, this is compensated by the time the model needed for completing its training, since it was 60% faster than the first one (8 hours and 20 hours respectively), saving significant amount of computation time.

## 5 Conclusion and Future Work

In this work, a classification model capable of recognizing 10 different categories of clothing images was presented. The process followed for analyzing the Deep Learning architectures of VGG, ResNet and a variation of ResNet were described in detail, as well as the techniques performed to find the optimal model and boost its performance.

DeepFashion was used for model training, while iMaterialist was used for evaluating the transferability of the produced model. The work was mainly focused on hyperparameter tuning, which is a necessary but time-consuming process

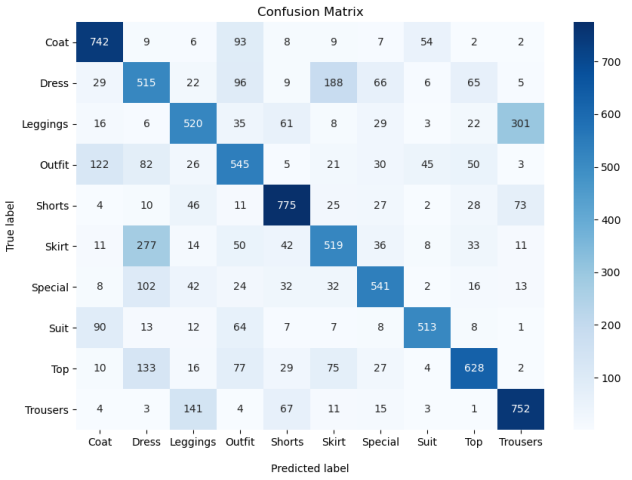


**Figure 3.** Retrained ResNet50 on iMaterialist based on the pretrained weights

for achieving the highest accuracy. The produced model achieved 80.5% accuracy on DeepFashion, while the fine-tuning of the pre-trained model on iMaterialist led to an 62.5% accuracy with a 60% reduction in training time, compared to the corresponding model trained from scratch.

Future work involves the improvement of the input datasets by manually refining its misplaced labels, which can be precisely identified using already trained models and even its enhancement with more samples, in order for the produced model to provide more robust results. Moreover, a wider set of experiments can be conducted in order to improve the performance of the model, such as further investigation on selecting a proper model architecture, detailed tuning of the hyperparameters in the pre-trained model's fine-tuning process and testing other training techniques in the fine-tuning process.

## Acknowledgments

This research has been co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T1EDK-03464)

## References

[1] Kuan-Ting Chen and Jiebo Luo. 2016. When Fashion Meets Big Data: Discriminative Mining of Best Selling Clothing Features. *CoRR* abs/1611.03915 (2016). arXiv:1611.03915 http://arxiv.org/abs/1611.03915

[2] J. Cychnerski, A. Brzeski, A. Boguszewski, M. Marmolowski, and M. Trojanowicz. 2017. Clothes detection and classification using convolutional neural networks. In *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. 1–8.

[3] Yuying Ge, Ruimao Zhang, Lingyun Wu, Xiaogang Wang, Xiaoou Tang, and Ping Luo. 2019. DeepFashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images. *CoRR* abs/1901.07973 (2019). arXiv:1901.07973 http://arxiv.org/abs/1901.07973

[4] Xavier Glorot and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (Proceedings of Machine Learning Research, Vol. 9)*, Yee Whye Teh and Mike Titterington (Eds.). PMLR, Chia Laguna Resort, Sardinia, Italy, 249–256. http://proceedings.mlr.press/v9/glorot10a.html

[5] Congying Guan, Sheng-feng Qin, Wessie Ling, and Guofu Ding. 2016. Apparel recommendation system evolution: an empirical review. *International Journal of Clothing Science and Technology* 28 (11 2016), 854–879. https://doi.org/10.1108/IJCST-09-2015-0100

[6] Sheng Guo, Weilin Huang, Xiao Zhang, Prasanna Srikhanta, Yin Cui, Yuan Li, Matthew R. Scott, Hartwig Adam, and Serge J. Belongie. 2019. The iMaterialist Fashion Attribute Dataset. *CoRR* abs/1906.05750 (2019). arXiv:1906.05750 http://arxiv.org/abs/1906.05750

[7] Z.X. Guo, W. Wong, SYS Leung, and Min Li. 2011. Applications of artificial intelligence in the apparel industry: A review. *Textile Research Journal* 81 (11 2011), 1871–1892. https://doi.org/10.1177/0040517511411968

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*. 1026–1034.

[9] Shintami Chusnul Hidayati, Chuang-Wen You, Wen-Huang Cheng, and Kai-Lung Hua. 2018. Learning and Recognition of Clothing Genres From Full-Body Images. *IEEE Transactions on Cybernetics* 48 (2018), 1647–1659.

[10] Riaz Ullah Khan, Xiaosong Zhang, Rajesh Kumar, and Emelia Opoku Aboagye. 2018. Evaluating the Performance of ResNet Model Based on Image Recognition. In *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence* (Chengdu, China) *(ICCAI 2018)*. Association for Computing Machinery, New York, NY, USA, 86–90. https://doi.org/10.1145/3194452.3194461

[11] Maria Th Kotouza, Sotirios-Filippos Tsarouchis, Alexandros-Charalampos Kyprianidis, Antonios C Chrysopoulos, and Pericles A Mitkas. 2020. Towards Fashion Recommendation: An AI System for Clothing Data Retrieval and Analysis. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 433–444.

[12] Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, and Xiaoou Tang. 2016. DeepFashion: Powering Robust Clothes Recognition and Retrieval With Rich Annotations. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[13] E.W.T. Ngai, S. Peng, Paul Alexander, and Karen Moon. 2014. Decision support and intelligent systems in the textile and apparel supply chain: An academic review of research articles. *Expert Systems with Applications: An International Journal* 41 (01 2014), 81–91. https://doi.org/10.1016/j.eswa.2013.07.013

[14] Karen Simonyan and Andrew Zisserman. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 [cs.CV]

[15] S. Vittayakorn, K. Yamaguchi, A. C. Berg, and T. L. Berg. 2015. Runway to Realway: Visual Analysis of Fashion. In *2015 IEEE Winter Conference on Applications of Computer Vision*. 951–958.