

# Derin Modellerin Eğitiminde Uygulanan Yöntemlerin Giysi Niteliklerinin Tespiti Üzerinde Kıyaslamalı Bir Analizi

## A Comparative Analysis of Practices in Training Deep Models for Fashion Attribute Detection

Mustafa Sercan Amaç, Aykut Erdem ve Erkut Erdem  
Bilgisayar Mühendisliği  
Hacettepe Üniversitesi  
Ankara, Türkiye  
sercanamac@gmail.com ,(aykut,erkut)@cs.hacettepe.edu.tr

**Özetçe** —Akıllı telefon teknolojilerinin ve sosyal medya uygulamalarının hızla artmasıyla birlikte, her gün milyarlarca fotoğrafın insanlar tarafından kişisel cihazlarıyla paylaşıldığı bir zamanda yaşamaktayız ve bu fotoğrafların büyük bir kısmı kişi görüntülerini veya özçekimleri içermekte. Bu çalışmada, kişi görüntülerinde giysi niteliklerini tanıma ve sınıflandırma problemini incelemekteyiz. Yakın zaman önce bu amaçla toplanan, büyük ölçekli bir veri kümesi olan StreetStyle-27k veri kümesi üzerinde gerçekleştirdiğimiz kapsamlı deneylerde derin evrimsel ağların etkili eğitimi için kullanılan en güncel yöntemlerden olan İlk yeniden başlatmalı olasılıksal bayır inisi, Odak Kayıp, Sıcaklıkla ölçekleme yöntemlerinin bir analizini gerçekleştirmekteyiz. Özellikle, bizimki gibi vahşi doğada tabir edilen problemlerde sıklıkla gözlemlenen bir sorun olan etiketlerin dağılımlarının dengesiz olduğu durumda öğrenmeyi ele almaktayız. Elde ettiğimiz en iyi sonuçlar StreetStyle-27k veri kümesi çalışmasında elde edilen sonuçlardan %3.67 daha yüksektir. Bu çalışmanın diğer araştırmacıların çalışmalarına ışık tutacağı umuyoruz.

**Anahtar Kelimeler**—giysi nitelikleri, derin öğrenme, çoklu görev öğrenimi

**Abstract**—With the rapid increase of smartphone technologies and social media apps, we live in a time where every day billions of photographs are shared by people through their personal devices, and a large amount of these photos involves person images or selfies. In this study, we investigate the problem of recognizing and classifying fashion attributes in person images. We perform extensive experiments on the StreetStyle-27k dataset with the, a recently proposed large-scale dataset collected for this purpose, in which we analyze the current best practices such as Stochastic Gradient Descent with Warm Restarts, Focal loss, Temperature Scaling that are generally used for effective training of deep convolutional networks. Especially, we elaborate on a specific challenge that commonly arise in in-the-wild problems such as ours, which is learning when the distribution of labels is unbalanced. The results we get with the best model is %3.67 better than StreetStyle. We hope that our results will shed some light and be useful to other researchers.

**Keywords**—fashion attributes, deep learning, multi task learning

### I. GİRİŞ

Akıllı telefonların ve tabletlerin günlük hayatın ayrılmaz birer parçası haline gelmesiyle beraber bu cihazlar üzerinden sosyal medyada her gün milyarlarca yazı ve görsel materyal paylaşılmaktadır. Yüklenen bu görsel veri içinde miktar olarak en büyük yoğunluğu kişi görüntüleri veya özçekimler (*selfies*) oluşturmaktadır. Bu kaynaklardan toplanan görüntüler üzerinde giysilerin görsel analizinin gerçekleştirilmesi son yıllarda bilgisayarla görme literatüründe dikkat çeken bir konu haline gelmiştir. Bu çalışmada biz bu amaçla toplanan veri kümelerinden biri olan StreetStyle27k üzerinde yaptığımız kişi resimleri üzerinden kıyafet etiketleme başarısı üzerine yapılan analiz ele alınacaktır.

### II. İLGİLİ ÇALIŞMALAR

İnsanların giydikleri kıyafetler, bunları giyen kişilerin nitelikleri de düşünülecek olursa görsel açıdan oldukça zengin çeşitlilikte örnekler içermektedir. Bu bağlamda giysilerin sınıflandırılması, analizi, vb. konular üzerine bir çok araştırma literatürde gerçekleştirilmiş durumdadır. Örneğin [2]'de, araştırmacılar kişilerin nitelikleri ve pozlarını tahmin etmeye yönelik bir yaklaşım önermişlerdir. Bir başka çalışma olan [6]'da ise yazarlar otomatik kıyafet kombini yapabilen bir yapay öğrenme sistemi geliştirmişlerdir. Benzer bir çalışmada [1], görüntülerden renk ve benzeri öznitelikler otomatik olarak çıkartılıp rasgele orman tekniğiyle kişi görüntüleri üzerinden giyim stillerinin sınıflandırılması çalışılmıştır. Araştırmacılar [5]'da yine bu yönde bir çalışma ortaya koymuşlar ve DeepFashion [8] veri kümesini kullanarak belirli giysi niteliklerini tahmin edebilen bir derin model sunmuşlardır.

### III. STREETSTYLE-27K VERİ KÜMESİ

StreetStyle-27K veri kümesi, sosyal ağlardan toparlanmış, insanların günlük hayatta çekilmiş fotoğraflarından oluşmaktadır (bkz. Şekil 1). Bu veri kümesindeki yer alan yaklaşık 27 bin fotoğraf, insanların giymekte oldukları giysilere özel 12 farklı nitelikte etiketlenmiştir. Ayrıca



Şekil 1: StreetStyle-27K veri kümesinden örnek görüntüler.

fotoğrafta yer alan kişinin üst vücut bölgesini kapsayan alanın koordinat bilgisi de veri kümesinde sağlanmış durumdadır. Bu çalışmamızda esas aldığımız giysi niteliklerinin kestirimleri, bu koordinatlar kullanılarak kırılan kişi görüntüleri üzerinde gerçekleştirilmektedir. Tablo 1, ele alınan giysi niteliklerini, bu niteliklere sahip örnek sayılarını göstermektedir. Buradan da gözlemlenebileceği üzere giysi nitelikleri, “Ceket Var” ve “Ceket Yok” gibi ikili sınıflar içerebildikleri gibi “Baskın Renk” niteliğinde olduğu gibi çoklu sınıflandırmalar da içerebilmektedirler. Bu veri kümesindeki en önemli sorun veriler arasındaki dengesizliklerdir. Örneğin, veri kümesindeki gözlük takmayan insan sayısı takan insan sayısının yaklaşık 10 katıdır. Bu yüzden, düz stratejiler kullanılarak eğitilen bir model her zaman hayır cevabını verse de 0.90 gibi yüksek bir başarı oranı yakalayabilmektedir. Veri kümesi aynı zamanda resmin çekildiği yer, saat gibi meta bilgilere de sahiptir ancak biz bu çalışmada yalnızca giysi nitelikleri üzerine bir analiz çalışması ortaya koymaktayız.

#### IV. YÖNTEM

Çalışmamızda giysi niteliklerinin kestirimi için derin öğrenme tabanlı modellerin başarımları üzerinde bir analiz ortaya konmaktadır. Baz aldığımız derin modeller çoklu görev öğrenim (*multi task learning*) tekniğine dayalı olarak eğitilen evrimsel sinir ağları (*convolutional neural networks*) modellerinden oluşmaktadır. Özellikle belirtmemiz gerekirse, kullandığımız modellerin omurgasını ImageNet üzerinde eğitilmiş olan Resnet50 [4] ve Inception V3 [12] modelleri oluşturmaktadır. Gerçekleştirdiğimiz analizlerde bu modeller StreetStyle-27K veri kümesinde ince ayarlamaya (*fine-tuning*) tabi tutularak giysi niteliklerini tahmin etmek üzere tekrar eğitilmektedir. Her bir nitelik için ayrı derin modeller eğitmek yerine parametrelerin paylaşılması prensibine dayanan çoklu öğrenme yapısı kullanılarak 12 farklı giysi niteliğini tek başına tahmin edebilecek modellerin eğitimine yoğunlaşmıştır. Bunu gerçekleştirmek için Resnet50 ve Inception V3 modellerinin tamamen bağlı olan (*fully-connected*) son katmanlarının ardından etiketlenmiş olan 12 farklı nitelik için sınıflandırma yapılmasına imkan veren yeni katmanlar tanımlanmış ve önceki katmanlarla birlikte eğitime tabi tutulmuştur. Çalışmamızda bu çoklu öğrenmeye dayanan modellerin eğitiminde dört farklı eğitim strateji test edilmiştir. Bunlar 1. Hızlanmalı Olasılıksal Bayır İnişi (*Stochastic Gradient Descent with Momentum- SGD*) [11], 2. İlk Yeniden

Tablo I: StreetStyle-27K veri kümesindeki giysi etiketlerinin dağılımı.

	Hayır	Evete	Baskın Renk	Giym Kategorisi	Yaka
Ceket	18078	7113	Siyah (6545)	Gömlek (4666)	Yuvarlak (9799)
Yaka	16774	7299	Beyaz (4461)	Dış giyim (4580)	Katlı (8119)
Atkı/Eşarp	23979	1452	2+ renk (2439)	T-shirt (4580)	V-yaka (2017)
Kravat	24843	827	Mavi (2419)	Elbise (2558)	
Şapka	23279	2255	Grı (1345)	Kolsuz (1348)	
Gözlük	22058	3401	Kırmızı (1131)	Takım Elbise (1143)	
Çok katlı	15921	8829	Pembe (649)	Kazak (874)	
			Yeşil (526)		
			Sarı (441)		
			Kahverengi (386)		
			Mor (170)		
			Turuncu (162)		
			CamGöbeği (33)		
				Kol Uzunluğu	
				Uzun Kol (13410)	
				Kısa Kol (7145)	
				Kolsuz (3520)	
					Giym Stili
					Düz (15933)
					Grafik (3832)
					Çizgili (1069)
					Çiçekli (885)
					Kareli (532)
					Noktalı (241)

Başlatmalı Olasılıksal Bayır İnişi (*Stochastic Gradient Descent with Warm Restart- SDGWR*) [9], 3. Odaksal Kayıp Fonksiyonu (*Focal Loss*) [7] ve 4. Sıcaklık Ölçekleme (*Temperature Scaling*) [3] yaklaşımlarıdır. Aşağıdaki altbölümlerde bu yaklaşımların detayları verilmektedir.

#### A. Hızlanmalı Olasılıksal Bayır İnişi

Derin modellerin öğrenilmesi gereken çok fazla sayıda parametreye sahip karmaşık modellerdir. Bu parametreler gözetimli öğrenme problemleri özelinde etiketli, büyük çaplı veri örnekleri üzerinden tanımlı bir kayıp fonksiyonun eniyelenmesi ile bulunmaktadır. Bu modellerin eğitimde sıklıkla kullanılan yaklaşımlardan biri olan momentumlu olasılıksal bayır inişi yaklaşımında bilinmeyen model parametreleri aşağıdaki yinelemeli denklem dizisi ile hesaplanmaktadır:

$$\mathbf{v}_{t+1} = \mu_t \mathbf{v}_t - \eta_t \nabla_{f_t}(\mathbf{x}_t) \quad (1)$$

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{v}_{t+1} \quad (2)$$

Burada  $t$  yapay zaman değişkenini,  $\mathbf{x}_t$   $t$  anındaki model parametrelerini,  $\mathbf{v}_t$  başlangıçta 0'a ayarlanmış hız vektörünü,  $f_t$   $t$  anındaki örnek yığını için hesaplanan kayıp fonksiyonu değerini,  $\mu_t$  hızlanma oranını, ve  $\eta_t$  öğrenme hızını ifade etmektedir. Deneylelerimizde giysilerin niteliklerini kestiren derin modeller  $\eta_0 = 0.01$ , momentum değeri olarak 0.9, ağırlık azalımı olarak 0.0001 hiperparametreleri kullanılarak tasdik kaybı (*validation loss*) artmaya başlayıya kadar eğitilmiştir.

#### B. İlk Yeniden Başlatmalı Olasılıksal Bayır İnişi

Derin öğrenme modellerini eğitirken modelin lokal optimalara takılması oldukça sık rastlanan bir durumdur. Bu bağlamda parametre uzayında daha iyi başarımları işaret eden optima noktaları bulmak için önerilen yaklaşımlardan biri de önceki bölümde ele aldığımız olasılıksal bayır inişi yaklaşımı için yeniden başlatma stratejilerinin kullanılmasıdır. Bu stratejilerden biri olan, [9]'de önerilen ılık yeniden başlatmalı SGD yaklaşımında öğrenme hızlı, belirli aralıklarla bazı değerlerle başlatılmakta ve planlı olarak azaltılmaktadır. Bu yeniden başlatma ve değer atamaları  $i$ 'nci çalıştırma için aşağıdaki denkleme göre yapılmaktadır:

$$\eta_t = \eta_{min}^i + \frac{1}{2} (\eta_{max}^i - \eta_{min}^i) \left( 1 + \cos \left( \frac{T_{cur}}{T_i} \pi \right) \right) \quad (3)$$

Burada  $T_{cur}$  son yeniden başlatmadan sonra geçen dönem (*epoch*) sayısını ifade etmektedir.  $T_i$  ise yeniden başlatmanın gerçekleştirileceği dönem sayısına karşılık gelmektedir ve bu

sayı her yeniden başlatmada  $T_{mul}$  ile ifade edilen katsayı ile arttırılmaktadır:

$$T_i = T_{cur} + T_i * T_{mul} \quad (4)$$

Şekil 2’de olasılıksal bayır inişi için ılık yeniden başlatma kullanılan bir öğrenme hızı planı gösterilmiştir. Deneylerimizde giysilerin niteliklerini kestiren derin modeller  $\eta_0 = 0.01$ ,  $T_{mul} = 3$ ,  $T_i = 1000$  ve momentum değeri olarak 0.9 hiperparametreleri kullanılarak tasdik kaybı (*validation loss*) artmaya başlayıyaya kadar eğitilmiştir.

### C. Odak Kayıp Fonksiyonu

Odak kayıp fonksiyonu, literatürde özellikle tek aşamalı bir yol izleyen derin nesne tanıma yaklaşımları için önerilmiş olan ve özellikle verideki sınıf dengesizliğin kaynaklı öğrenme zorluklarını aşmak için kullanılan bir kayıp fonksiyonudur. Kısaca özetlemek gerekirse, bu kayıp fonksiyonu, çapraz düzensizlik kayıp fonksiyonun aşağıda belirtilen şekilde yeniden biçimlendirilmiş bir halidir:

$$FL(p_t) = -(1 - p_t)^\gamma \log(p_t) \quad (5)$$

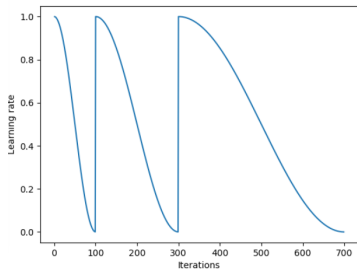
Burada,  $p$  derin modelin hesapladığı sınıf olasılığını ifade ettiği varsayılırsa  $p_t$  de şu şekilde tanımlanmaz:

$$p_t = \begin{cases} p & \text{eğer } y = 1 \\ 1 - p & \text{aksi halde} \end{cases} \quad (6)$$

Denklem 5’te  $(1 - p_t)^\gamma$  terimi standard çapraz düzensizlik kayıp fonksiyonunu modüle ederek, kolay öğrenilen sınıflardan gelen kayıpları düşürmekte ve zor öğrenilen ve örnek sayısı az olan sınıflara daha çok önem verilmesini sağlamaktadır. Deneylerimiz odak kayıp fonksiyonu kullanılarak gerçekleştirilen eğitimlerde  $\gamma$  parametresi değeri 2 olarak alınmıştır.

### D. Sıcaklık Ölçekleme

Günümüzde önerilmiş olan derin modellerde sınıflar için doğru doğruluk olabilirliliğini temsil eden olasılık hesaplarının kestirimi önemli bir problemdir. Bu amaçta literatürde işlem sonrası adımı olarak izlenen yollardan biri tahminler üzerinde kalibrasyon yapılmasıdır. Çalışmamızda bu tarz bir yöntem olan sıcaklık ölçekleme yöntemi [3] giysi niteliklerini tespit eden derin modellerin kalibrasyonunda kullanılmıştır. Bu yaklaşım, tüm sınıflar için sıcaklığı ifade eden tek bir skalar



Şekil 2: Öğrenme hızını ayarlamak için kullanılabilecek bir yeniden başlatma planı. Bu plan için parametreler  $T_i = 100$ ,  $\eta_0 = 1$ , ve  $T_{mul} = 2$  olarak alınmıştır.

parametre,  $T > 0$ , kullanarak yeni güven tahminlerini aşağıdaki denklemle hesaplamaktadır:

$$q_i = \max_k \sigma_{SM}(\mathbf{z}_i/T)^{(k)} \quad (7)$$

Burada  $\mathbf{z}_i$  logit vektörünü,  $\sigma_{SM}$  eşiksiz en büyük fonksiyonunu (*softmax function*) ve  $k$   $K$  çoklu sınıftan birini ifade etmektedir. Bu çalışmada ele alınan problem bağlamında 12 farklı giysi niteliği tahmin edildiğinden dolayı sıcaklık ölçekleme yaklaşımıyla hesaplanan 12 ayrı logit vektörünün her biri için bir sıcaklık değeri hesaplanmıştır. Bu sıcaklık değerleri bulunurken de eniyileme yöntemi olarak Adam ve kayıp fonksiyonu olarak da çapraz düzensizlik fonksiyonu kullanılmıştır.

Gerçekleştirilen kalibrasyonun etkisini ölçmek için kullanılan empirik ölçütlerden biri beklenen kalibrasyon hatasıdır (*expected calibration error* - ECE). Bu hata ölçütü şu şekilde tanımlanmaktadır:

$$ECE = \sum_{m=1}^M \frac{|B_m|}{n} |\text{acc}(B_m) - \text{conf}(B_m)| \quad (8)$$

Bu ölçüt hesaplanırken tahminler eşit aralıklı  $M$  kutucuğa bölünmekte ve bu kutucukların doğruluk (*accuracy* - acc) ve güven (*confidence* - conf) farklarının ağırlıklı ortalaması hesaplanmaktadır. Ortalamada kullanılan ağırlıklar,  $|B_m|$  ile temsil edilen bir kutucuğa düşen örnek sayısının  $n$  ile gösterilen toplam örnek sayısına oranıyla hesaplanmaktadır. Deneylerimizde ECE azaldıkça modelin örnek sayısı az olan etiketlerde daha iyi tahmin yapabildiği gözlemlenmiştir.

## V. DENEYLER

Tablo II’de Resnet50 ve Inception V3 modellerini omurga olarak alan, StreetStyle-27K veri kümesinin eğitim örnekleri ile çoklu öğrenme kullanılarak eğitilen modellerin başarımları, [10]’da rapor edilmiş sonuçlarla birlikte gösterilmektedir. Bu tablolardan görülebileceği üzere çalışmamızda baz aldığımız her iki derin model omurgası için ilgili eğitim sürecinde ılık yeniden başlatmalı olasılıksal bayır inişi yaklaşımı kullanıldığında modellerin nitelik kestirim başarımlarında artış gözlemlenmiştir. Odak kayıp fonksiyonunun eğitimde kullanılması hem tahmin oranını iyileştirmiş, hem de verideki dengesizlikten kaynaklanan yanlış pozitifleri büyük ölçüde ortadan kaldırmıştır. Bu bağlamda odak kaybı ile eğitilen modellerde ayrıca kalibrasyona ihtiyacı olmadığı gözlemlenmiştir. Ancak odak kayıp fonksiyonunun eğitim için kullanılmadığı durumlar için sıcaklık ölçeklemeye bağlı olarak gerçekleştirilen model kalibrasyonu veri dengesizliğini dengelemede başarılı olmuştur. Şekil 3’de veri kümesinden örnek bir görüntü üzerinde aynı modelin kalibre edilmeden önce ve edildikten sonraki halleri arasındaki tahmin farkı gösterilmektedir. Uygulanan kalibrasyonun net bir şekilde veri dengesizliğinden (veri kümesinde baskın rengin sarı ve şapka giyilen görüntü sayısının diğerlerine nazaran çok daha az gözlemlenmesinden) kaynaklanan başarısızlığı azalttığı gözlemlenebilmektedir.

Deneylerimizde genel olarak en başarılı tahmin sonuçları Inception v3 omurgası üzerinde SGDWR ve Odak Kayıp stratejileri birlikte kullanılarak eğitim yapıldığında yakalanmıştır ve bu başarımların değerleri StreetStyle çalışmasında [10] rapor edilmiş sonuçlardan %3.67 daha yüksektir. Şekil 4’te eğitimde kullanılmaması bazı örnek

Tablo II: Giysi niteliklerinin kestirimlerine ait model performansları

Strateji	Giyim Stili	Baskın Renk	Kravat	Yaka	Atkı/Eşarp	Kol Uzunluğu	Yaka Şekli	Giyim Kategorisi	Ceket	Şapka	Gözlük	Çok Katlı	Ortalama	
ResNet50	SGD	%83.91	%63.55	%98.12	%86.78	%96.12	%74.33	%82.13	%64.96	%86.15	%94.41	%95.20	%84.31	%84.16
	SGDWR	%85.66	%70.80	%98.30	%86.95	%96.43	%78.11	%83.19	%68.13	%87.06	%96.38	%98.45	%84.39	%86.16
	Odak Kayıp	%87.07	%72.07	%98.23	%88.14	%96.14	%79.81	%83.40	%66.17	%88.16	%96.33	%97.53	%84.65	%86.48
	SGDWR + Odak Kayıp	%89.44	%76.73	%98.88	%89.73	%96.70	%82.81	%84.92	%72.24	%89.19	%97.05	%98.68	%87.04	%88.62
	SGDWR + Kalibrasyon	%86.02	%70.79	%98.15	%87.92	%96.49	%77.94	%82.55	%66.48	%87.72	%96.28	%98.57	%84.41	%86.11
Inception V3	SGD	%85.44	%66.79	%98.75	%87.36	%96.74	%75.05	%81.93	%65.42	%86.67	%95.52	%98.31	%84.98	%85.25
	SGDWR	%86.88	%70.57	%99.09	%89.42	%96.51	%80.96	%83.42	%70.09	%89.02	%96.64	%98.49	%85.81	%87.24
	Odak Kayıp	%89.77	%77.10	%99.10	%89.54	%96.51	%82.85	%85.11	%72.40	%89.98	%97.01	%98.49	%86.54	%88.69
	SGDWR + Odak Kayıp	%89.85	%76.79	%98.94	%89.85	%97.52	%83.55	%86.4	%72.04	%90.66	%97.04	%98.79	%88.53	%89.16
	SGDWR + Kalibrasyon	%87.61	%71.35	%98.94	%90.08	%96.48	%80.91	%84.71	%69.70	%88.42	%96.82	%98.56	%87.06	%87.55
StreetStyle [10]	%85.30	%68.80	%97.90	%86.90	%94.40	%79.40	%83.10	%66.10	%86.90	%95.90	%98.20	%83.00	%85.49	



	Orjinal	Kalibrasyonlu
Baskın Renk	2+ Renk	Sarı
Şapka	Hayır	Evet
Ceket	Evet	Evet
Yaka	Hayır	Hayır
Kravat	Hayır	Hayır
Gözlük	Hayır	Hayır
Çok katlı	Evet	Evet
Giyim Kategorisi	Dışarı Giyisi	Dışarı Giyisi
Giyim Stili	Düz	Düz
Yaka Şekli	Katlı Yaka	Katlı Yaka
Kol Uzunluğu	Uzun	Uzun
Atkı/Eşarp	Hayır	Hayır

Şekil 3: Kalibrasyonun diğer etiketleri etkilemeden sonuçları iyileştirdiğini gösteren bir örnek.



Düz giyim,  
mavi renk,  
gözlüklü,  
uzun kol



Düz giyim,  
siyah,  
kravat,  
yaka,  
uzun kollu,  
takım elbise,  
ceketli,  
katlı yaka



Kareli,  
2+ renk,  
yaka,  
gözlüklü,  
katlı yaka,  
uzun kol



Noktalı,  
Gri,  
Kazak,  
Çok katlı,  
uzun kollu,  
yuvarlak yaka

Şekil 4: En başarılı modelimizin örnek giysi nitelik tahminleri.

görüntüler için bu modelimizin yüksek güvenle tahmin ettiği nitelikler gösterilmektedir.

## VI. SONUÇ

Bu çalışmada kişi niteliklerinin tespiti problemi üzerinden derin modellerin eğitiminde kullanılan farklı yaklaşımların detaylı bir analizi gerçekleştirilmiştir. Sosyal medyada paylaşılmış kişi fotoğrafları içeren StreetStyle-27K veri kümesi üzerinde elde ettiğimiz sonuçlar, çalışmamızda analiz edilen yaklaşımların genel olarak başarılı kestirimler yapabildiklerini göstermiştir. Denemiş olduğumuz SGDWR tekniği, standard SGD stratejine kıyasla derin kestirim modelimizi daha iyi bir optimaya ulaştırmıştır ve onun kestirim başarımını iyileştirmiştir. Benzer şekilde sıcaklık ölçekleme yönteminin nöral ağırları kalibre etmede başarılı olduğu gözlemlenmiştir. Ayrıca yapılan deneyler sonucunda fokal kayıp fonksiyonun eğitime dahil edilmesinin üzerinde çalıştığımız veride mevcut olan nitelik dengesizlikleriyle başa çıkabilmeyi sağlamış ve başarımı pozitif yönde etkilemiştir. Devam çalışmamızda, StreetStyle-27K veri kümesinden ilhamla Türkiye'ye ve Türk insanının moda eğilimlerine özel yeni bir veri kümesi toparlamayı ve buradakin benzer analizleri bu veri üzerinde gerçekleştirmeyi planlamaktayız.

## TEŞEKKÜR

Bu çalışma TÜBİTAK (116E685 nolu 1003 projesi) ve Türkiye Bilimler Akademisi (TÜBA-GEBİP 2018 programı) tarafından desteklenmiştir.

## KAYNAKÇA

- [1] L. Bossard, M. Dantone, C. Leistner, C. Wengert, T. Quack, and L. Van Gool. Apparel classification with style. In *ACCV*, pages 321–335. Springer, 2012.
- [2] H. Chen, A. Gallagher, and B. Girod. Describing clothing by semantic attributes. In *ECCV*, pages 609–623. Springer, 2012.
- [3] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger. On calibration of modern neural networks. In *ICML*, 2017.
- [4] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, volume 6, pages 770–778.
- [5] M. Jia, Y. Zhou, M. Shi, and B. Hariharan. A deep-learning-based fashion attributes detection model. *CoRR*, abs/1810.10148, 2018.
- [6] Y. Li, L. Cao, J. Zhu, and J. Luo. Mining fashion outfit composition using an end-to-end deep learning approach on set data. *IEEE Transactions on Multimedia*, 19(8):1946–1955, 2017.
- [7] T.-Y. Lin, P. Goyal, R. B. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. In *ICCV*, pages 2999–3007, 2017.
- [8] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang. DeepFashion: Powering robust clothes recognition and retrieval with rich annotations. In *CVPR*, pages 1096–1104, 2016.
- [9] I. Loshchilov and F. Hutter. SGDR: Stochastic gradient descent with warm restarts, 2017.
- [10] K. Matzen, K. Bala, and N. Snavely. Streetstyle: Exploring world-wide clothing styles from millions of photos. *CoRR*, abs/1706.01869, 2017.
- [11] I. Sutskever, J. Martens, G. Dahl, and G. Hinton. On the importance of initialization and momentum in deep learning. In *ICML*, 2013.
- [12] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *CVPR*, 2016.