

İnsan Görsel Karmaşıklık Kararlarının Derin Öğrenme ile Tahmin Edilmesi

Predicting Human Visual Complexity Judgments via Deep Learning

Hamit Başgöl¹, Emre Uğur²

¹Bilişsel Bilim, Boğaziçi Üniversitesi, İstanbul, Türkiye

²Bilgisayar Mühendisliği, Boğaziçi Üniversitesi, İstanbul, Türkiye

Özetçe —Görsel karmaşıklık bir görselin tanımlanmasındaki güçlük veya bir görselin içerisinde yer alan detayların fazlalığı olarak tanımlanabilir. Görsel karmaşıklığın çalışılması hem insan görsel sistemini anlamak hem de insan kullanımına uygun ürünler üretmek için önemlidir. Literatürde, görsel karmaşıklığı ölçmek için geliştirilen yöntemler mevcuttur. Bu çalışmada derin öğrenme transfer ile öğrenme yöntemi bağlamında kullanılarak görselleri karmaşıklıkları açısından birbiriyle karşılaştırabilen ve insan performansı ile korelasyon gösteren iki model geliştirilmiştir. Modeller hangi görselin diğerinden daha karmaşık bulunduğunu yüksek başarıyla tahmin edebilmektedir. Kategoriler-içi model için %95 ve kategoriler-arası model için %78 doğruluk oranı elde edilmiştir. Bununla birlikte, modellerin veri setindeki karmaşıklık değerleri ile göstermiş olduğu korelasyonlar alandaki diğer algoritmaların korelasyonlarıyla karşılaştırılmıştır. Karşılaştırma sonucunda eğitilen modellerin korelasyonlarının alandaki algoritmaların korelasyonlarından yüksek olduğu tespit edilmiştir. En yüksek korelasyonu gösteren model dikkate alındığında; sahneler için 0.77, reklamlar için 0.81, görselleştirmeler için 0.84, objeler için 0.77, iç tasarım için 0.89, sanat için 0.85, süpermatizm için 0.84 ve bütün görseller için 0.78 korelasyon katsayısına ulaşılmıştır.

Anahtar Kelimeler—görsel karmaşıklık, derin öğrenme, transfer öğrenimi, bilişsel bilim, psikoloji

Abstract—Visual complexity can be defined as the difficulty of defining an image or the number of details in the image. Studying visual complexity is important for understanding human visual system and developing systems for human use. In the literature, there are studies trying to measure visual complexity of images. In this study, two systems that are capable of comparing two images in terms of their visual complexities and show correlation with human visual complexity judgments are developed by using deep learning in the context of transfer learning. Models can predict which image is found more complex than the other with a substantial success. 95% accuracy score was obtained for the between-domain and 78% was obtained for the cross-domain model. Additionally, correlations of models with complexity values in the dataset were compared with correlations of algorithms in the literature. As a result of the comparison, it was determined that correlations of models were higher than correlations of algorithms. Considering the best correlations; we achieved 0.77 for scenes, 0.81 for advertisements, 0.84 for visualizations, 0.77 for objects, 0.89 for interior design, 0.85 for

art, 0.84 for suprematism, 0.78 for all images.

Keywords—visual complexity, deep learning, transfer learning, cognitive science, psychology

I. GİRİŞ

Görme yetisi bize hareket etme, objeleri kullanma ve çevrede olanları algılama şansı vermektedir. Görme yetisinin oluşumu evrimsel tarihimizdeki çok erken dönemlere rastlamaktadır [1]. Görsel karmaşıklık bir görselin tanımlanmasındaki güçlük veya bir görselin içerisinde yer alan görsel detayların karışıklığı ve fazlalığı olarak tanımlanabilir [2], [3]. Psikolojideki uğraşılardan biri görsel karmaşıklığı anlamaktır [4]. Görsel karmaşıklık uygulamalı açıdan da anlamlıdır; bunun nedeni görsel karmaşıklık ile estetik arasındaki bağlantıdır. İdeal uyarılma teorisine göre bir uyarının tercih edilmesi ve uyarının karmaşıklığı arasında U biçiminde bir ilişki mevcuttur. Orta düzeyde karmaşık bir uyarın çok veya az karmaşık uyarılara kıyasla tercih edilir [5], [6].

Görsel karmaşıklığın simetri ile negatif [7] ve görsellerdeki çizgi ve iç açı sayısı ile pozitif bir ilişkisinin olduğu tespit edilmiştir. [8], [9]. Görsellerdeki yapısal elemanlardan yola çıkarak geliştirilen bir çalışmada, elemanları otomatik olarak sayan ve karmaşıklık değerine yakınsamaya çalışan bir model geliştirilmiştir [10]. Görsellerin sıkıştırılabilirliğinin içlerindeki bilgi yoğunluğunu dolayısıyla görsel karmaşıklık değerlerini vereceğini savunan çalışmalar mevcuttur [11]. Bu çalışmalar algoritmik bilgi teorisine bağlı olarak, karmaşıklık değerinin bir resmi temsil etmek için kullanılan minimum kodla elde edilebileceğini savunur [4], [11]. Bu görüş ZIP ve JPEG sıkıştırma formatlarından sonra ortaya çıkan dosya boyutunun görsel karmaşıklık, hata ve arama zamanıyla ilişkisinin bulunması sayesinde desteklenmiştir [11].

Bir görselde temsil edilen sahneye ve görselin içerisinde yer alan objelere yönelik bilginin de karmaşıklık değerlerini etkilemesi beklenmektedir. Fakat bu bilgiler hesaplamalı modellerde dikkate alınmamıştır. Literatürde yer alan algoritmalar kenar yoğunluğu, öznelik yoğunluğu, altbant entropisi [12], sıkıştırma oranı [13] ve alan sayısıdır [14] [15]. İnsanlar obje

bilgisini gerektiren karmaşıklık değerlerinin tespitini kolaylıkla gerçekleştirse de, söz konusu algoritmaların performanslarının bu konuda zayıf olduğu [15] tarafından tespit edilmiştir.

Derin öğrenme, obje tanıma [16], [17], video gruplandırma [18] ve el yazısı tanıma [19] gibi çalışmalarda sıkça kullanılmaktadır; fakat derin öğrenmenin görsel karmaşıklık literatürünü etkilediği söylenemez. Literatürde yapay sinir ağları kullanılan ve görsel karmaşıklıkla ilişkili iki çalışma tespit edilebilmiştir. İlk çalışmada, otokodlayıcı ile insanların estetik kararları modellenmiştir [20]. Bir diğer çalışmada ise görsel karmaşıklık değerlerini tahmin edebilmek için öznetelik mühendisliği yardımıyla yapay sinir ağı geliştirilmiştir [21]. Fakat çalışmada sanatsal ve soyut resimler kullanılmıştır. Derin öğrenme sayesinde bu problemlerin önüne geçilebileceği düşünülmekte ama veri miktarının düşük olması sebebiyle bunun zor olduğu daha önce literatürde belirtilmiştir [22]. Veri sayısının düşüklüğünün aşırı öğrenmeye yol açtığı bilinmektedir. Transfer öğrenimi daha önce başka bir görev için eğitilen sistemlerin yeni görevler için kullanılabilmesidir [23]. Transfer öğrenimi ile daha az veri kullanılarak sistemler eğitilebilir [24]. Bu nedenle, transfer ile öğrenme bu problemde kullanılabilir. Transfer öğrenimi kullanmanın yanında, veri setinde karmaşıklık değerleri olarak yer alan sayıları öğrenmek yerine (regresyon), modeller iki farklı görselin hangisinin daha karmaşık olduğunun bulunması (kategorizasyon) için eğitilebilir. Bunun veri miktarındaki artışa yol açacağı ve karmaşıklığı oluşturan temel görsel bilgilerin öğrenilmesine neden olacağı düşünülmektedir.

Bu çalışmanın amacı, derin öğrenmeyi transfer ile öğrenme bağlamında kullanarak insanların görsel karmaşıklık hakkındaki kararlarını tahmin edecek bir model geliştirmektir.

II. YÖNTEM

A. Veri Seti

Çalışmada SAVOIAS olarak adlandırılan ve toplamda 7 farklı kategoriye ayrılmış 1420 görsel içeren veri seti kullanılmıştır [15]. Veri setinde yer alan kategoriler sırasıyla sahneler (200 görsel), reklamlar (200 görsel), görselleştirmeler ve bilgi grafikleri (200 görsel), objeler (200 görsel), iç tasarım (100 görsel), sanat (420 görsel) ve süprematizm (100 görsel)'dir. Bu veri seti görsel karmaşıklık çalışmalarında kullanılan veri setlerindeki yaygın problemlerin giderilmesi için geliştirilmiştir. Bu problemlerden biri çalışmalarda kullanılan resimlerin paylaşılmasındır. Bir diğeri görsellerin içeriklerindeki, görünüşlerindeki ve karmaşıklıklarının derecelendirilmesindeki tutarsızlıklardır. Veri setini hazırlayan araştırmacılar toplam 1687 katılımcıdan 37.000 görsel çifti için görsel karmaşıklık karşılaştırmaları yapmalarını istemişler ve bu karşılaştırmaları 0 ile 100 arasındaki karmaşıklık değerlerine çevirmişlerdir. Veri seti oluşturulduktan sonra literatürdeki en iyi algoritmalar ile test edilmiş ve her bir kategori için çeşitli skorlar çıkarılmıştır. Resimlerin karmaşıklık değerleri ve karşılık gelen frekanslar Şekil 1'de görülebilir. Ayrıca veri setindeki kategorilere karşılık gelecek görseller bilgi vermek açısından örneklendirilmiştir. Bu görseller Şekil 2'de görülebilir.

B. Veri İşleme

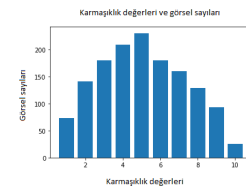
Modeller, veri setinde yer alan görsellerin karmaşıklık değerlerini tahmin etmek yerine (regresyon), verilen iki görselden

hangisinin daha karmaşık olduğunu tahmin etmek üzere (kategorizasyon) eğitilmiştir. Bu nedenle, daha önce bahsedilen veri seti [15] yardımıyla yeni bir veri seti oluşturulmuştur. Bu veri setini oluşturabilmek için öncelikle kategoriler, %70'i eğitim setini (994 görsel) ve %30'u test setini (426 görsel) oluşturmak üzere ikiye bölünmüştür. Kategoriler-içi modeli eğitebilmek için görseller kategoriler içi karşılaştırılmış; kategoriler-arası modeli eğitebilmek için kategorilerden bağımsız olarak her görsel bir diğeriyle karşılaştırılmıştır. Karşılaştırma sırasında eğer verilen ilk görsel diğeriinden daha karmaşık veya diğeriyle eşit karmaşıklık değerine sahipse bu verinin kategorisi 1; eğer verilen ikinci görsel birinci resimden daha karmaşık ise bu verinin kategorisi 2 olarak belirlenmiştir. Bunun nedeni eşit karmaşıklık değerinde olan resimlerin görece az olması, dolaşısıyla eşitliğin ayrı bir kategori olarak kullanılamamasıdır.

C. Model

Transfer öğreniminde kullanılacak temel model olarak MobileNetV2 seçilmiştir [21]. Bu modelin seçilmesinin nedeni modelin mobil cihazlar üzerinde çalışmak için geliştirilmesi ve buna bağlı olarak hesaplama açısından verimli olmasıdır. MobileNetV2, ImageNet ile yapılan çalışmalarda %74.7 obje tanıma skoruna ulaşmıştır [21]. MobileNetV2'nin sonunda yer alan ve objelerin kategorizasyonunda rol oynayan katman çıkarılmıştır. Bu katman çıkarıldıktan sonra, MobileNetV2, bu çalışmada eğitilen derin öğrenme modeline verilecek iki görselden öznetelikler oluşturmak amacıyla paylaşımlı olarak kullanılmıştır. Üzerinde değişiklik yapılan MobileNetV2'den görseller için elde edilen aktivasyonlar vektör haline getirilmiş ve daha sonra yığın normalizasyonu (batch normalization) işlemine tabi tutulmuştur. Bu işleme sokulduktan sonra görselleri temsil eden her bir vektör, lineer aktivasyon fonksiyonuna sahip 4 nörondan oluşan katmanlardan geçirildikten sonra uç uca eklenmiştir. Daha sonra bu vektör iki nörona sahip bir çıktı katmanına bağlanmış ve çıktı katmanı için softmax fonksiyonu belirlenmiştir. Bu nöronun amacı eriştiği vektör yardımıyla hangi görselin daha karmaşık olduğunu belirten kategorileri oluşturmaktır. Modelin girdi biçimi her görsel için 224 x 224 x 3 olarak belirlenmiştir. Bu yapı iki farklı biçimde eğitilerek kategoriler-içi ve kategoriler-arası olarak adlandırılan iki model oluşturulmuştur.

Kategoriler-içi model yalnızca aynı kategorideki görsellerin kendi içinde karşılaştırılmasıyla ve yine aynı kategorideki görsellerin test setinde kullanılmasıyla oluşturulmuştur. Kategoriler-arası modelde kategori farkı gözletilmemiş ve modelden farklı kategorilerdeki görselleri de karşılaştırması beklenmiştir. İki farklı modelin eğitilmesinin temel amacı görsel karmaşıklık kıyaslamalarında kategori bilgisinin etkisini tespit edebilmektir. Modellerin eğitim bilgileri ve performansları Tablo I'de verilmiştir. Yapılan denemelerde kategoriler-arası modelin test setindeki performansının daha fazla miktarda veri kullanılsa dahi artmadığı gözlemlenmiştir. Bu nedenle



Şekil 1. Veri setinde yer alan görsellerin karmaşıklıkları ve görsel sayıları



Şekil 2. Veri setinde yer alan kategoriler ve örnek görseller

kategoriler-arası modelin veri miktarı kategoriler-içi modellerle benzer tutulmuştur. Kategoriler-arası modelin kategoriler-içi modele kıyasla aşırı öğrenmeye daha yatkın olduğu ve artan veri miktarının bu trendi olumlu yönde etkilemediği tespit edilmiştir.

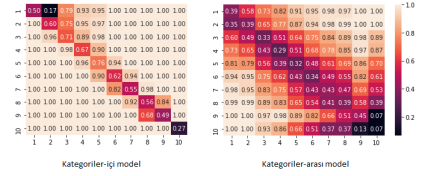
III. ANALİZLER

Eğitim sonucunda erişilen performans bilgileri Tablo I'de bulunabilir. Kategoriler-içi modelin kategoriler-arası modelden daha iyi performans gösterdiği görülmektedir. Modellerin doğruluk skorlarıyla birlikte her bir model için ayrıca Pearson korelasyon katsayısı hesaplanmıştır. Korelasyon katsayılarını hesaplamak için test setinde yer alan görseller eğitim setinde yer alan görsellerle karşılaştırılmıştır. Eğer test setindeki bir görsel eğitim setindeki bir görselden daha karmaşık bulundysa, test setindeki görselin karmaşıklık skoruna 1 eklenmiştir ve böylece test setindeki her görsel için bir karmaşıklık skoru elde edilmiştir. Daha sonra elde edilen bu skorlar test seti için belirlenmiş gerçek karmaşıklık değerleriyle karşılaştırılmıştır. Her iki modelin karmaşıklık değerlerinin, gerçek karmaşıklık değerleri ile korelasyonuna Tablo II'den ulaşılabilir. Kategoriler-içi modelin kategoriler-arası modele kıyasla karmaşıklık skorlarıyla olan ilişkisi daha yüksektir. Fakat, kategoriler-arası modelin sahneler ve süprematizm kategorilerinde daha yüksek korelasyona ulaşması ilgi çekicidir. Bunun nedeni Şekil 4'ten de anlaşılabilceği üzere modelin Süprematizm kategorisini aşırı öğrenmesi olabilir.

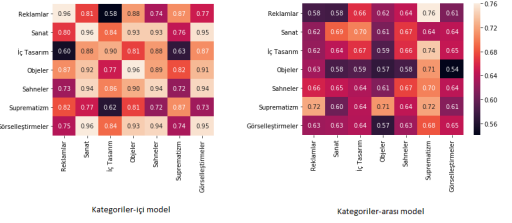
Eğitilen her bir model için karmaşıklık skorları 10 grup oluşturacak şekilde kategorize edilerek doğruluk matrisleri üretilmiştir. Doğruluk matrisleri Şekil 3'de görülebilir. Kategoriler-içi model için üretilen doğruluk matrisi dikkate alındığında, modelin hangi görselin daha karmaşık olduğunu doğru şekilde verebildiği gözlenmektedir. Görsellerin karmaşıklık değerleri birbirine çok yaklaştığında modelin kararlarındaki doğruluk azalmaktadır. Aynı durum kategoriler-arası model için daha şiddetli bir biçimde görülmektedir. Modelin verdiği kararların doğru olması için karmaşıklık skorları

TABLO I. MODEL EĞİTİM VE PERFORMANS BİLGİLERİ

Bilgiler	Eğitilen Modeller	
	Kategoriler-içi Model	Kategoriler-arası Model
Eğitim seti veri sayısı	128304	145180
Test seti veri sayısı	32076	36295
Toplam veri sayısı	160380	181475
Epok sayısı	7	1
Optimizasyon fonksiyonu	Adam	Adam
Yığın sayısı	64	64
Öğrenme oranı	0.001	0.001
Doğruluk skoru	%95	%78
Hassasiyet skoru (0 - 1)	0.96 - 0.95	0.80 - 0.76
Geri çağırma skoru (0 - 1)	0.95 - 0.95	0.75 - 0.80
F1 skoru (0 - 1)	0.96 - 0.95	0.77 - 0.78



Şekil 3. Kategoriler-arası ve kategoriler-içi modeller için oluşturulan doğruluk matrisleri



Şekil 4. Kategoriler-arası ve kategoriler-içi modellerin kategori-bazlı doğruluk matrisleri

arasında ciddi bir fark olması gerekmektedir. Kategoriler-arası ve kategoriler-içi modelden alınan doğruluk skorları Şekil 4'de verilmiştir. Kategoriler-içi modeli test edebilmek için kategoriler arası karşılaştırma yapılacak şekilde yeni bir test seti oluşturulmuştur. İlk dikkat çeken bulgu kategoriler-arası modelin performansının aynı kategorideki karşılaştırmalarda dahi çok düşük olması hatta farklı kategorilerdeki karşılaştırmaların kimi zaman çok daha iyi sonuç vermesidir. Bununla beraber, kategoriler-içi modelin performansının aynı kategorideki görseller için yapılan karşılaştırmalarda %90'ın üzerinde olduğu görülmektedir. Ayrıca kategoriler-içi model, farklı kategorilerin karşılaştırılmasında -bu şekilde eğitilmiş olmasa dahi- kategoriler-arası modelden daha iyi performans göstermektedir.

Geliştirilen modellerin ve alandaki gelişmiş algoritmaların veri setinde yer alan görseller ile korelasyonu Tablo II'te verilmiştir [15]. Her iki modelin diğer algoritmalarla kıyasla çok daha iyi performans gösterdiği görülmektedir. Özellikle diğer algoritmaların kötü performans gösterdiği sahneler ve objeler kategorilerinde modelin çok iyi performans göstermesi, MobileNetV2 modelinin daha önce öğrendiği bilgileri bu görev için kullanması nedeniyle olabilir. Kategoriler-içi modelin de en düşük performans gösterdiği kategorilerin 0.77 korelasyon ile sahneler ve objeler kategorileri olduğu unutulmamalıdır. Yalnızca alan sayısı algoritması süprematizm kategorisinde kategoriler-içi model ile karşılaştırılabilir performans göstermiştir. Her ne kadar kategoriler-içi modelin performansı diğer algoritmaların performanslarından daha iyi olsa da, bu model objeler ve sahneler kategorilerinde aynı hata yanlılığını göstermiştir. Kategoriler-arası modelin performansı genel olarak kategoriler-içi modelden düşüktür; fakat diğer algoritmalarla kıyasla daha iyi olduğu görülmektedir. Şekil 5'te kategoriler-içi modele verilen iki görsel ve belirlenlik haritaları görülebilir. Mavi ton yoğunluğu olan bölgelerin model tarafından daha karmaşık olarak tanımlandığı tahmin edilmektedir.

IV. TARTIŞMA

Bu çalışmada görsel karmaşıklık kararlarını tahmin edilecek transfer öğrenimini temel alan iki derin öğrenme

TABLO II. MODELLERİN DİĞER ALGORİTMALARLA KARŞILAŞTIRILMASI

Kategoriler	Korelasyon Değerleri						
	Kenar Yoğunluğu	Sıkıştırma Oranı	Alan Sayısı	Öznitelik Yoğunluğu	Altbant Entropisi	Kategoriler-içi Model	Kategoriler-arası Model
Sahneler	0.16	0.30	0.57	0.42	0.16	0.77	0.83
Reklamlar	0.54	0.56	0.41	0.56	0.54	0.81	0.69
Görselleştirmeler	0.57	0.55	0.38	0.52	0.61	0.84	0.72
Objeler	0.28	0.16	0.29	0.30	0.10	0.77	0.68
İç Tasarım	0.61	0.68	0.67	0.58	0.31	0.89	0.83
Sanat	0.48	0.51	0.65	0.22	0.33	0.85	0.80
Süpremizm	0.18	0.60	0.84	0.48	0.39	0.84	0.90
Hepsi	-	-	-	-	-	0.78	0.74



Şekil 5. Modele verilen görseller ve model tarafından oluşturulan belirginlik haritaları. İlk görsel (kedi) ikinci resimden (iguana) daha karmaşık bulunmuştur.

modeli geliştirilmiştir. Hem aynı kategorideki resimlerin birbirleriyle karşılaştırılmasıyla elde edilen kategoriler-içi, hem de farklı kategorilerin karşılaştırılmasıyla eğitilen kategoriler-arası model alandaki algoritmalarla kıyasla daha iyi performans göstermiştir. Fakat kategoriler-arası model kategoriler-içi modele kıyasla daha zayıftır. Bunun nedeninin insanların farklı kategorilerdeki görseller için farklı özelliklere dikkat etmesi olduğu söylenebilir. Bu çalışmanın eksikliklerinden biri transfer öğreniminde yalnızca bir modelin kullanılmasıdır. Gelecekteki çalışmalar birden fazla modelin ayrı ayrı eğitilmesi veya birden fazla modelden alınan özelliklerin birleştirilerek eğitilmesiyle daha iyi sonuçlara ulaşabilir. Ayrıca bu çalışmada eğitilen modellerin gösterdiği kadarıyla kategoriler arasında yapılan karşılaştırmalarda performans düşüşü yaşanabilmektedir.

Görsel bilginin verimli bir biçimde kullanılması insan gündelik yaşamı için çok önemlidir. Bu nedenle insan kullanımına uygun ürünler geliştirmek için hangi görsel uyarıların insanlar için daha tercih edilir olduğunu anlamak önemlidir. Sonuç olarak, görsel karmaşıklığın anlaşılması için yapılacak çalışmalar önem arz etmektedir.

KAYNAKLAR

- [1] W. J. Gehring, "The evolution of vision," *Wiley Interdisciplinary Reviews: Developmental Biology*, vol. 3, no. 1, pp. 1–40, 2014.
- [2] A. Forsythe, "Visual complexity: is that all there is?" in *International Conference on Engineering Psychology and Cognitive Ergonomics*. Springer, 2009, pp. 158–166.
- [3] C. Heaps and S. Handel, "Similarity and features of natural textures," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 25, no. 2, p. 299, 1999.
- [4] D. C. Donderi, "Visual complexity: a review," *Psychological bulletin*, vol. 132, no. 1, p. 73, 2006.
- [5] D. E. Berlyne, "Novelty, complexity, and hedonic value," *Perception & Psychophysics*, vol. 8, no. 5, pp. 279–286, 1970.
- [6] A. Forsythe, M. Nadal, N. Sheehy, C. J. Cela-Conde, and M. Sawey, "Predicting beauty: fractal dimension and visual complexity in art," *British journal of psychology*, vol. 102, no. 1, pp. 49–70, 2011.
- [7] A. Gartus and H. Leder, "The small step toward asymmetry: Aesthetic judgment of broken symmetries," *i-Perception*, vol. 4, no. 5, pp. 361–364, 2013.

- [8] J. Hochberg and V. Brooks, "The psychophysics of form: Reversible-perspective drawings of spatial objects," *The American Journal of Psychology*, vol. 73, no. 3, pp. 337–354, 1960.
- [9] M. García, A. N. Badre, and J. T. Stasko, "Development and validation of icons varying in their abstractness," *Interacting with computers*, vol. 6, no. 2, pp. 191–211, 1994.
- [10] A. Forsythe, N. Sheehy, and M. Sawey, "Measuring icon complexity: An automated analysis," *Behavior Research Methods, Instruments, & Computers*, vol. 35, no. 2, pp. 334–342, 2003.
- [11] D. C. Donderi and S. McFadden, "Compressed file length predicts search time and errors on visual displays," *Displays*, vol. 26, no. 2, pp. 71–78, 2005.
- [12] R. Rosenholtz, Y. Li, and L. Nakano, "Measuring visual clutter," *Journal of vision*, vol. 7, no. 2, pp. 17–17, 2007.
- [13] S. Corchs, F. Gasparini, and R. Schettini, "No reference image quality classification for jpeg-distorted images," *Digital Signal Processing*, vol. 30, pp. 86–100, 2014.
- [14] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 5, pp. 603–619, 2002.
- [15] E. Saraee, M. Jalal, and M. Betke, "Savoiias: A diverse, multi-category visual complexity dataset," *arXiv preprint arXiv:1810.01771*, 2018.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [17] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [18] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2014, pp. 1725–1732.
- [19] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner *et al.*, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [20] Y. Tan, Y. Zhou, G. Li, and A. Huang, "Computational aesthetics of photos quality assessment based on improved artificial neural network combined with an autoencoder technique," *Neurocomputing*, vol. 188, pp. 50–62, 2016.
- [21] P. Machado, J. Romero, M. Nadal, A. Santos, J. Correia, and A. Carballal, "Computerized measures of visual complexity," *Acta psychologica*, vol. 160, pp. 43–57, 2015.
- [22] C. Fernandez-Lozano, A. Carballal, P. Machado, A. Santos, and J. Romero, "Visual complexity modelling based on image features fusion of multiple kernels," *PeerJ*, vol. 7, p. e7075, 2019.
- [23] S. Pan and Q. Yang, "A survey on transfer learning. IEEE transaction on knowledge discovery and data engineering, 22 (10)," 2010.
- [24] W. Zhao, "Research on the deep learning of the small sample data based on transfer learning," in *AIP Conference Proceedings*, vol. 1864, no. 1. AIP Publishing, 2017, p. 020018.