

Sosyal Robotlar için Çocuklarda Yüz Eylem Birimi Tespitinde Maskeleye ve Sıkıştırma Teknikleri

Masking and Compression Techniques for Efficient Action Unit Detection of Children for Social Robots

Erhan Biçer*, Şeyma Takır*, Cemal Gürpınar*, Pınar Uluer*[†], Hatice Köse*

*Yapay Zeka ve Veri Mühendisliği Bölümü,

İstanbul Teknik Üniversitesi, İstanbul, Türkiye

E-Posta: {bicer21, takir21, gurpinarcemal, uluerp, hatice.kose}@itu.edu.tr

[†]Bilgisayar Mühendisliği Bölümü,

Galatasaray Üniversitesi, İstanbul, Türkiye, E-Posta: puluer@gsu.edu.tr

Özetçe—Bu çalışmada çocuklar için geliştirilen bir yardımcı sosyal robot sisteminin duygu tanıma modülünde kullanılmak üzere önerilen yüz verisi maskeleye ve sıkıştırma yöntemleri sunulmaktadır. Çalışmanın amacı robotun üzerinde sınırlı işlem gücü ve hafıza kullanımına dayanan ve gerçek zamanlı bir eylem birimi tabanlı duygu tanıma modülü oluşturmaktır. Bu amaçla çeşitli ön işleme ve görüntü sıkıştırma yöntemlerinin çocuk yüz verilerinde eylem birimi tanıma etkisi gözlemlenmiştir. Sonuçlar, gerçek zamanlı uygulamalarımızda robotun tepki süresini kısıtlayacak işlemlerin hariç tutulabileceğini göstermektedir.

Anahtar Kelimeler—yüz eylem birimi tespiti, resim maskeleye, görüntü sıkıştırma, insan-robot etkileşimi.

Abstract—This paper proposes masking and compression techniques for action unit detection of children to be used on a social robot. The aim of the study is to develop a real-time action unit-based emotion recognition module based on limited processing power and lower memory consumption on the robot. For this purpose, the effect of various preprocessing techniques and image compression methods on action unit recognition of children are observed. The results show that preprocessing steps that will constrain the robot's response time, can be excluded in our real-time applications.

Keywords—facial action unit detection, image masking, image compression, human-robot interaction.

I. GİRİŞ

Son yıllarda yardımcı sosyal robot uygulamalarında artış bulunmaktadır. Engelli ve yaşlı bireylerin günlük hayatlarına yardımcı sosyal robot uygulamalarının yanısıra, çocukların eğitim hayatlarını destekleyici [1] ve engelli bireylerin sosyal becerilerini arttırmak için kullanılan sosyal robotlar geliştirilmektedir [2]. Bu kullanıcı gruplarının davranış ve duygu durumlarının tanınması robotların kullanıcıya daha uygun tepki verebilmesine ve etkileşimlerinin daha verimli olmasına imkan sağlar. Özellikle otizmli çocuklar gibi özel ihtiyaçları olan çocukların terapi sürecinde kullanılan sosyal robotların çocukların sosyalleşmesi ve diğer insanlarla iletişimi açısından önemli bir rol oynadığı, robotun çocukların duygularını anlamasının etkileşim kalitesini iyileştireceği ve robotun kabul edilebilirliğini arttıracak ön sürülmektedir [3].

Bu çalışmada bu amaçla EMBOA projesinin (<https://emboa.eu/>) bir parçası olarak otizmli çocukların yüz ifadesinden duygu tanımayı sağlayan ve bir sosyal robot üzerinde gerçek zamanlı çalışabilecek bir modül geliştirilmesi planlanmıştır. Çalışmanın ilerde işitme engelli çocuklar için de kullanılması planlanmaktadır. Gerçek zamanlı robot uygulamaları için geliştirilen yöntemler arasında görüntü sıkıştırma gibi tekniklere yer verilmektedir [4]. Bu tekniklerin yanı sıra, hafif ağırlıklı ve verimli derin sinir ağı modelleri öneren çalışmalar da vardır [5]. Bizim çalışmamızda ise veri ön işleme aşamalarını daha verimli hale getirerek ve veri sıkıştırma algoritmalarını kullanarak bu probleme bir çözüm sunulmuştur.

II. MASKELEME

Yüz verisinden duygu ve eylem birimi tanıma çalışmalarında kullanılan görüntüler, resimler arasındaki farklılıkları azaltmak için ışıklandırma düzeltme, yüz tespiti, yüz hizalama, vb. gibi ön işleme adımlarından geçer. Çünkü ışıklandırma ve poz farklılıkları gibi farklılıklar eylem birimi tanımayı daha zorlu kılar [6].

Daha önce geliştirilen bir eylem birimi tanıma modelinde de benzer ön işleme teknikleri kullanılmıştır [7]: (1) Veri arttırma (döndürme, yatay çevirme, kaydırma), (2) Yüz tespiti, (3) Yüze göre kırpmaya işlemi, (4) Görüntüyü yüze göre hizalama, (5) Görüntüye sadece yüz görünecek şekilde bir maske uygulanması, (6) Işıklandırma düzeltme tekniği (Histogram Eşitleme veya CLAHE).

Bu çalışmada uygulanan ön işleme akışına göre 4. madde ve sonrasındaki işlemlerin tamamı "maskeleye" olarak adlandırılmıştır. Çalışmamızdaki ön işleme akışındaki yüz tespit işlemleri için DLib [8] ve OpenCV kütüphaneleri kullanılmıştır.

Yüzden gerçek zamanlı duygu tanıma uygulamalarında, modele sınıflandırması için bir yüz görüntüsü verilirken, yukarıda belirtilen ön işleme tekniklerinden 2. ve 3. adımlar uygulanacaktır. Ancak, zaman ve kaynak kullanımı açısından değerlendirildiğinde, sonrasındaki işlemlerin (4-6. adımlar) maliyetinin yüksek olduğu görülür. Bu durum, özellikle gerçek zamanlı uygulamalar açısından dezavantajlı olabilir. Bu

durumun analizi için “maskeleme” işleminin gerektirdiği ek süre ölçülmüştür. Söz konusu zaman ölçümü Nvidia RTX3070 Laptop GPU, Intel Core i7-11800H @2.30GHz, 16 GB RAM içeren bir dizüstü bilgisayarda yapılmıştır. 50 resim girdisinde maskeleme işlemi ortalama 1.0592 ± 0.0344 saniye, maskeleme işlemi olmadan, yalnızca yüz kırıldığında, ortalama 0.1209 ± 0.0152 saniye sürdüğü tespit edilmiştir.

Bu çalışma ile uyguladığımız ön işleme adımlarının eylem birimi tanıma doğruluğuna olan etkisini gözlemlemek amaçlanmıştır. Duygu ifadeleri sırasında yüzdeki kas aktivite değişikliği 250 ms ile 5 s arasında sürmektedir [9]. Gerçek zamanlı yüz eylem birimi tanıma sistemlerinin saniyede 0.4 ila 8 kare arası işlemesi beklenmektedir. Yaklaşık 1 saniye süren maskeleme işlemi, sistemin yavaşlamasına ve bazı duygu geçişlerinin kaybedilmesine neden olabilir.

A. Veri Seti

CAFE (Child Affective Facial Expression) [10], 2 ila 8 yaş arası çocukların çeşitli duygu ifadelerinin yer aldığı resimler içeren bir veri setidir. Bahse konu veri seti; üzgün, mutlu, şaşkın, kızgın, iğrenmiş, korkmuş ve nötr olmak üzere 7 ana duygu etiketi içerir. Bu çalışmada, CAFE veri setindeki çocukların resimleri bir uzman tarafından FACS’e [11] uygun olarak 18 eylem birimi ile tekrar etiketlenmiştir. Etiketlenen bu eylem birimlerinin verideki gözlemlenme sıklıklarına göre en sık gözlemlenen altısı (AU1, AU4, AU6, AU9, AU12 ve AU17) bu çalışmada kullanılmıştır.

CAFE veri setindeki resimlere ön işleme uygulanarak iki farklı sürüm oluşturulmuştur: maskeli ve maskesiz. Maskeli CAFE veri setine bir önceki bölümde bahsedilen 6 ön işleme tekniğinin tümü uygulanmıştır, maskesiz veri setine ise bunlardan sadece ilk üçü uygulanmıştır. Şekil 1’de bu işlemler ChildEFES veri setinden [12] seçilen örnek bir resim üzerinde gösterilmiştir.



Şekil 1: Ham veri (en sol), maskesiz (orta) ve maskeli (en sağ)

B. Işıklandırma Düzeltme Teknikleri

Yüz resimlerindeki ışıklandırma farkı eylem birimi analizini zorlu kılan faktörlerden biridir [6]. Buradan yola çıkarak, maskeleme aşamasında resimlere histogram eşitleme tekniği uygulanmıştır. Bu teknik, resimdeki yoğunluk değerlerini dağıtarak kontrastı iyileştirir [13], [14].

Işıklandırma düzeltme tekniğinin maskeleme aşamasındaki performans artışına etkisinin analizi için yapılan deneylerde maskeli ve maskesiz veri setlerine ışıklandırma düzeltme tekniği uygulanıp modelin tanıma performansı gözlemlenmiştir. Histogram eşitlemeye ek olarak, CLAHE tekniği de kullanılmıştır. CLAHE histogram eşitlemeyi genişleterek, resimdeki eşit bölünmüş alt alanlara histogram eşitleme uygulayarak histogramın yüksekliği belirlenen eşik değerini aşmayacak

şekilde histogram tekrar dağıtılır [15]. Bu iki teknik, Şekil 2’de görselleştirilmiştir.



Şekil 2: Histogram eşitleme (orta) ve CLAHE (en sağ)

C. Model Mimarisi

CAFE veri seti üzerinde veri arttırma uyguladıktan sonra toplamda 2344 resim oluşturulmuştur. Veri sayısının azlığından dolayı transfer öğrenme tercih edilmiştir. Bu kapsamda, VGG-16 [16] model mimarisi, önceden yüz tanıma için eğitilmiş VGG-Face [17] ağırlıkları ile kullanılmıştır. En son katman olarak düzleştirme (flatten) katmanı ile birlikte 6 nöronlu bir tam bağlı (fully connected) sınıflandırıcı katman eklenmiştir. Transfer öğrenme uygulanırken, veri sayısının yetersizliğinden ve veri sayısına göre yüksek parametre sayısının olmasından dolayı sınıflandırıcı katman hariç bütün katmanların ağırlıkları eğitime kapatılmıştır [18]. Çocuklar resimlerde birden fazla eylem birimi gösterebileceğinden sınıflandırma görevi çoklu etiketli bir sınıflandırmadır. Bu nedenle, sınıflandırıcı katmanda aktivasyon fonksiyonu olarak sigmoid ve kayıp fonksiyonu olarak da ikili çapraz entropi (binary cross-entropy) kullanılmıştır.

Kapsamlı bir karşılaştırma amacıyla 3-katlı çapraz doğrulama uygulanmıştır. Her bir kat için 15 farklı hiperparametre kombinasyonu kullanarak Weights and Biases aracı ile model eğitilmiştir [19]. Bu kombinasyonlarda optimizasyon algoritması olarak; Adam, RMSProp ve SGD(Stochastic Gradient Descent), yığın boyutu olarak 32, öğrenme katsayısı olarak; 0.01, 0.001, 0.0001, momentum olarak 0.9, ağırlık azaltma katsayısı olarak 0.001 ve 0.0001 kullanılmıştır. Adam ve 0.01 öğrenme katsayısı en iyi ortalama F1 veren parametreler olmuştur.

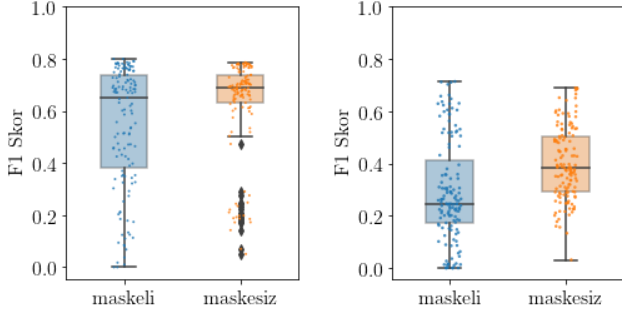
Bu 3-katlı çapraz doğrulama tekniği maskeli ve maskesiz verilerin her bir sürümü (histogram eşitleme uygulanmış, CLAHE uygulanmış, ve ışıklandırma düzeltme tekniği bulunmayan sürümler) için uygulanmış ve F1 sonuçları hesaplanmıştır. Olası aşırı öğrenme (overfitting) durumunun önüne geçebilmek adına eğer 5 adım (epoch) boyunca doğrulama setinde F1 sonucu artmaz ise öğrenme katsayısı 0.1 ile çarpılarak küçültülmüştür. Ayrıca, modelin doğrulama setindeki en iyi F1 performansının bulunduğu adımdaki ağırlıkları kaydedilmiş ve değerlendirmeler için modelin bu aşamadaki hali kullanılmıştır.

D. Değerlendirme Sonuçları

Değerlendirmelerde iki çeşit yöntem kullanılmıştır: (1) Öz değerlendirme: Maskeli ve maskesiz veri seti kendi veri setlerinden derlenen test setinde test edilmiştir, (2) Çapraz değerlendirme: Maskeli veri seti ile eğitilen modeller maskesiz veri seti ile, maskesiz veri seti ile eğitilen modeller maskeli veri seti ile test edilmiştir. Modellerin eğitildiği ve test edildiği veri setleri, kullanılan farklı ışıklandırma teknikleri ve elde edilen F1 skorlar TABLO I’de gösterilmiştir.

TABLO I: Maskeli ve Maskesiz Veri Setleri için F1 Skorları (Satırlarda en iyi sonucu verenler italik, iki deney grubu arasındaki en iyi skorlar sütunlarda koyu renkli gösterilmiştir.)

Eğitim Verisi	Test Verisi	F1 (Hist. Eş.)	F1 (CLAHE)	F1 (Yok)
Maskeli	Maskeli	63.97±21	64.91±18	34.48±17
Maskesiz	Maskesiz	62.58±19	63.59±17	62.51±19
Maskeli	Maskesiz	31.79±19	35.95±17	20.08±19
Maskesiz	Maskeli	43.62±14	43.33±13	32.78±11



Şekil 3: Öz (sol) ve Çapraz (sağ) Değerlendirme Sonuçları

Sonuçlar, modellerin eğitildiği veri seti ile test edildiğinde F1 skorlarının daha yüksek olduğunu göstermektedir. Ayrıca, maskeli veri ile eğitilen modellerin, maskesiz ile eğitilen modeller arasında istatistiksel olarak önemli bir fark olmadığı görülmüş, hatta ışıklandırma tekniği kullanılmadığında başarısında düşüş olduğu saptanmıştır. Bu durum, maskeleme işleminin yarattığı zamansal yükün eylem birimi tanıma açısından önemli bir getirisinin olmadığını göstermektedir. Dolayısıyla, gerçek zamanlı robotik uygulamalarımızda, eylem birimi tanıma ön işleme adımlarından maskeleme aşamasının hariç tutulması uygun görülmüştür. Ayrıca, Şekil 3'te görüldüğü gibi, maskesiz veri setinin öz değerlendirme ve çapraz değerlendirme F1 skorlarının ortanca değerinin maskelilere oranla daha yüksek olduğu gözlemlenmiştir. Maskeli veri setinin kullanıldığı değerlendirmelerde ise, F1 skor dağılımının daha yaygın olduğu görülmüştür. Dolayısıyla, kullanılan çeşitli ışıklandırma düzeltme algoritmalarının ve hiperparametrelerin farklı etkilerine karşı maskesiz eylem birimi tanıma performansının daha az değişkenlik gösterdiği dikkat çekmektedir. Bu gözlem, maskeli veri seti ile eğitilen modeller için uygun hiperparametre araştırmasını bulmanın önemini vurgulamaktadır, aksi takdirde, yüz eylem birimi tanıma doğruluklarında beklenmeyen düşüşlere yol açabileceğini göstermiştir.

III. GÖRÜNTÜ KALİTESİ

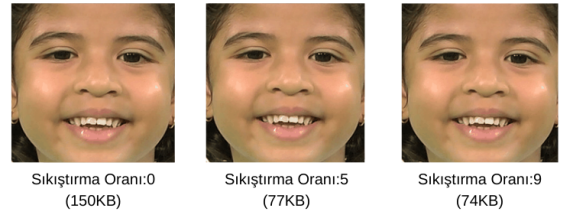
Çalışmanın bu kısmında CAFE veri seti kullanılarak farklı sıkıştırma tekniklerinin görüntü ve buna bağlı olarak eylem birimi tanıma performansı üzerindeki etkileri incelenmiştir. Hız ve depolama maliyetleri göz önünde bulundurularak farklı sıkıştırma teknikleri ve sıkıştırma oranları kullanılarak görüntü dosya boyutunun en aza indirilmesi amaçlanmıştır [20], [21].

Görüntü kalitesinin derin öğrenme modeli başarımına etkileri incelendiğinde görüntülere gürültü eklendiği, bulanıklaştırma (*blur*) filtresinden geçirildiği, JPEG veya JPEG2000 ile sıkıştırıldığı görülmüştür [22]. Bu çalışmada, görüntüler PNG formatında 3 farklı sıkıştırma oranı kullanılarak sıkıştırılmıştır. Daha sonra görüntüye dalgacık dönüşümü (*wavelet transform*)

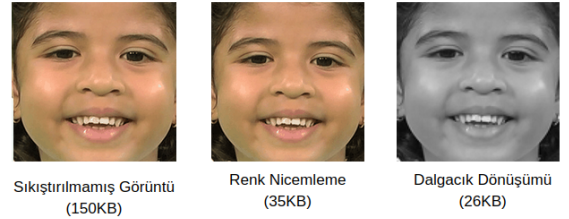
ve renk nicemleme (*color quantization*) metodları uygulanarak model başarısı gözlemlenmiştir.

A. Kayıpsız Sıkıştırma

Kayıpsız sıkıştırma, görüntüde herhangi bir kalite kaybı olmadan dosya boyutunun küçültülmesi işlemidir. Bu işlem sırasında dosyanın içinde işlevsiz olan meta verileri atılır. GIF, PNG, RAW gibi teknikler kayıpsız sıkıştırmayı desteklemektedir. Görüntüleri farklı oranlarda PNG formatında sıkıştırmak için birçok kütüphane bulunmaktadır [23]. Bu çalışmada kayıpsız sıkıştırma için LZM altyapısını kullanan PNG formatı seçilmiş ve PNG formatındaki görüntülerde PIL (Python Imaging Library) aracılığı ile sıkıştırma oranı olarak 0 (minimum), 5 (medyan) ve 9 (maksimum) değerleri kullanılmıştır. Şekil 4a'da gösterildiği gibi sıkıştırma oranı 0 verildiğinde - yani görüntü sıkıştırılmadığında - görüntü boyutunun 147KB olduğu görülmüştür. Sıkıştırma oranı 5 ve 9 verildiğinde ise görüntü boyutu yaklaşık 75KB olmuştur. Bu iki oran arasında dosya boyutundaki değişim açısından önemli bir fark gözlemlenmemiştir.



(a) Kayıpsız Sıkıştırılan Görüntüler



(b) Kayıplı Sıkıştırılan Görüntüler

Şekil 4: Kayıplı ve Kayıpsız Sıkıştırma

B. Kayıplı Sıkıştırma

Kayıplı sıkıştırma ile görüntü sıkıştırılırken bazı veriler kaybolur ve görüntü tekrar eski haline döndürülemez. Fakat bu teknik, görüntü boyutunu önemli ölçüde küçültür. Kayıplı sıkıştırma için bir çok farklı teknik mevcuttur. Bu çalışmada renk nicemleme ve dalgacık dönüşümü kullanılmıştır. Renk nicemleme yöntemi görüntüdeki renk sayısını azaltmaya dayanırken, dalgacık dönüşümü ise görüntüyü bulunduğu boyuttan daha farklı bir boyuta taşıyarak hesaplanan katsayılarla dayalı bir kodlama yapar. Görüntü elemanları arasındaki ilişkiyi en aza indirip sıkıştırma oranını arttırmayı hedefler. Bu yöntemin en önemli avantajlarından biri yüksek sıkıştırma oranları elde edilebilmesidir. Haar, coiflets, biorthogonal gibi farklı dalgacık dönüşümü teknikleri bulunmaktadır [23]. Bu çalışmada biorthogonal dalgacık dönüşümü kullanılmıştır. Şekil 4b'de görüldüğü gibi renk nicemleme ile sıkıştırma yapıldığında

görüntü boyutu yaklaşık 34KB olurken dalgacık dönüşümü ile 26KB civarında olmaktadır.

C. Değerlendirme Sonuçları

Dosya boyutunu küçültmek amacıyla veri setinde bulunan görüntüler farklı yöntemler kullanılarak sıkıştırılmıştır. Kayıpsız sıkıştırmada PNG formatı kullanırken kayıplı sıkıştırmada renk nicemleme ve dalgacık dönüşümü kullanılmıştır. Görüntüler sıkıştırılmadığında ortalama 150KB civarında olurken kayıpsız sıkıştırma ile bu boyut 74KB'a kadar düşmektedir. Kayıplı sıkıştırma ile dosya boyutunun 26KB'a kadar düştüğü görülmüştür.

Sıkıştırma yöntemlerinin model başarısı üzerindeki etkisini gözlemek için bu görüntüler kullanılarak VGG-Face modeli ile Bölüm II.A'da belirtilen 6 eylem birimi tespiti yapılmış ve sonuçlar daha sonra F1 skorlarına göre karşılaştırılmıştır. Elde edilen sonuçlara göre görüntüler kayıpsız sıkıştırıldığında sonuçlar birbirine çok yakın çıkmıştır: Görüntü sıkıştırılmadığında, yani PNG sıkıştırma oranı 0 seçildiğinde F1 skoru %71.29, bu oran 5 olduğunda %71.14, 9 olduğunda ise %71.26 çıkmıştır. F1 skorları arasında istatistiksel açıdan önemli bir fark görülmemektedir. Kayıplı sıkıştırmada ise, dalgacık dönüşümü için F1 skoru %70.90 iken renk nicemleme için bu skor %70.99'dur. Bu yöntemlerin F1 skorları arasındaki farkın da göz ardı edilebilir olduğu gözlemlenmiştir. Bu sonuçlar ışığında bu çalışma için optimum sıkıştırma tekniğinin dalgacık dönüşümü olduğu görülmüştür. Diğer tekniklerle karşılaştırıldığında hem dosya boyutu açısından avantajlı olduğu, hem de model performansına etkisi olmadığı gözlemlenmiştir.

IV. SONUÇ

Bu çalışma otizimli çocuklar (EMBOA Projesi) için geliştirilen bir yardımcı sosyal robot sisteminin duygu tanıma modülünde kullanılmak üzere önerilen yüz verisi maskeleme ve sıkıştırma yöntemlerini sunmaktadır. Bu amaçla eylem birimlerinin tanınmasında başarıyı düşürmeden yüksek işlem gücü ve hafıza kaynaklarına ihtiyaç duyan işlemlerin elenmesi için farklı yaklaşımlar denenmiştir. CAFE veri seti üzerinde yapılan çalışmaların sonuçlarına göre, maskesiz veriler ile eğitilen modellerin maskeli veriler ile eğitilen modellere kıyasla sonuçlarının yakın olduğu gözlemlenmiştir. Maskeleme ön işleminin gerçek zamanlı robotik uygulamalarda oluşturacağı olası zaman kaybı ve model performansı açısından önemli bir geliştirme sağlamadığından gelecek çalışmalara dahil edilmesi uygun görülmüştür. Farklı sıkıştırma teknikleri uygulandığında ise dalgacık dönüşüm tekniğinin en düşük görüntü dosya boyutunu elde etmeyi sağladığı gözlemlenmiştir. Bu çalışma kısıtlı kaynaklara ve işlemci gücüne sahip robotlarda gerçek zamanlı olarak kullanılacak yüksek başarılı duygu tanıma sistemlerinin geliştirilmesi ve yardımcı sosyal robotlarla etkileşimin daha verimli olması için de yararlı bir adım olacaktır.

TEŞEKKÜR

Bu çalışmada kullanılan Titan V ekran kartı ve Jetson TX2 modülü için NVIDIA'ya teşekkür ederiz. Çalışma Erasmus+ projesi EMBOA (Proje no: 2019-1-PL01-KA203-065096) tarafından desteklenmektedir.

KAYNAKLAR

- [1] D. Conti, S. D. Nuovo, S. Buono, and A. G. D. Nuovo, "Robots in education and care of children with developmental disabilities: A study on acceptance by experienced and future professionals," *International Journal of Social Robotics*, vol. 9, pp. 51–62, 2017.
- [2] N. Hurst, C. Clabaugh, R. Baynes, J. Cohn, D. Mitroff, and S. Scherer, "Social and emotional skills training with embodied moxie," 2020.
- [3] K. D. Bartl-Pokorny et al., "Robot-based intervention for children with autism spectrum disorder: a systematic literature review," *IEEE Access*, vol. 9, pp. 165 433–165 450, 2021.
- [4] O. Rippel and L. Bourdev, "Real-time adaptive image compression," 2017.
- [5] Y. Chen, H. Wu, T. Wang, Y. Wang, and Y. Liang, "Cross-modal representation learning for lightweight and accurate facial action unit detection," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 7619–7626, 2021.
- [6] R. Zhi, M. Liu, and D. Zhang, "A comprehensive survey on automatic facial action unit analysis," *The Visual Computer*, vol. 36, no. 5, pp. 1067–1093, May 2020.
- [7] E. Baglayici, C. Gurpinar, P. Uluer, and H. Kose, "A new facial expression processing system for an affectively aware robot," in *Pattern Recognition. ICPR International Workshops and Challenges*, A. Del Bimbo et al., Ed. Cham: Springer Int Publishing, 2021, pp. 36–51.
- [8] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.
- [9] B. Fasel and J. Luetttin, "Automatic facial expression analysis: a survey," *Pattern Recognition*, vol. 36, no. 1, pp. 259–275, 2003.
- [10] V. LoBue and C. Thrasher, "The child affective facial expression (cafe) set: Validity and reliability from untrained adults," *Frontiers in psychology*, vol. 5, p. 1532, 2015.
- [11] P. E. Ekman and E. L. Rosenberg, *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, 1997.
- [12] J. G. Negrão et al., "The child emotion facial expression set: A database for emotion recognition in children," *Frontiers in Psychology*, vol. 12, 2021.
- [13] S. Raj, S. Kumar, and S. Raj, "An improved histogram equalization technique for image contrast enhancement," in *International Conference on Communication and Computing*, 2015, pp. 1–6.
- [14] G. Yadav, S. Maheshwari, and A. Agarwal, "Contrast limited adaptive histogram equalization based enhancement for real time video system," in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2014, pp. 2392–2397.
- [15] G. Benitez-Garcia et al., "Face identification based on contrast limited adaptive histogram equalization (clahe)," in *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (ICIPV)*. Citeseer, 2011, p. 1.
- [16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations*, 2015.
- [17] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proceedings of the British Machine Vision Conference (BMVC)*, X. Xie, M. W. Jones, and G. K. L. Tam, Eds. Swansea, UK: BMVA Press, September 2015, pp. 41.1–41.12.
- [18] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Adv. Neural Inf. Process. Syst.*, Z. G. et al., Ed., vol. 27. Curran Associates, Inc., 2014.
- [19] L. Biewald, "Experiment tracking with weights and biases," 2020, software available from wandb.com.
- [20] M. Sindhu and R. Rajkamal, "Images and its compression techniques – a review," *International Journal of recent trends in engineering*, vol. 2, pp. 71–75, 11 2009.
- [21] A. Kaur and D. J. Singh, "Review on image compression techniques and advantages of image compression," *IJARSE Aug*, 2016.
- [22] S. Dodge and L. Karam, "Understanding how image quality affects deep neural networks," 2016.
- [23] R. Kaur and P. Choudhary, "A review of image compression techniques," *Int. J. Comput. Appl.*, vol. 142, pp. 8–11, 05 2016.