



Durağan Görüntüler Üzerinde Duygu Tanıma

Transfer Öğrenme ve Topluluk Öğrenmesi Kullanarak

Hüseyin ABANOZ, Zehra Çataltepe

İstanbul Teknik Üniversitesi, Bilgisayar Mühendisliği

1. Duygu Tanıma
2. Veri Kümesi
3. Transfer Öğrenme
4. Önceden Eğitilmiş Modeller
5. Eğitim
6. Deney Sonuçları
7. Karmaşıklığı Azaltma
8. Topluluk Öğrenmesi

Duygu Tanıma

Duygu Tanıma

Kızgınlık



Korku



İğrenme



Mutluluk



Üzüntü



Şaşkınlık

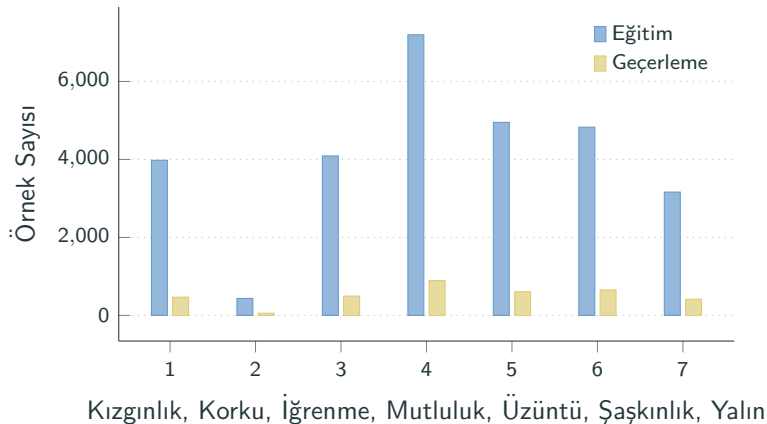


Yüz görüntüsünün Kızgınlık, İğrenme, Mutluluk, Üzüntü, Şaşkınlık ve duygu içermeyen Yalın sınıflarından biriyle etiketlenmesidir.

Veri Kümesi

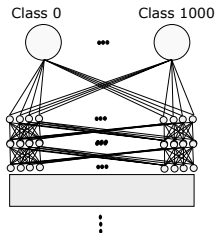
FER13 Kümesi:

- 32K gri renkli görüntü
- 48X48 piksel
- Ortalanmış Yüz Görüntüleri
- Kaynak internet, Crowdsourcing
- Sık kullanılma [1, 2, 3, 4, 5]

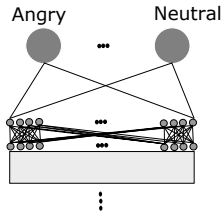


Transfer Öğrenme

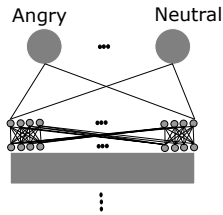
Transfer Öğrenme



Pre-Trained Network



Short Retraining



Long Retraining

● ■ Retrained ○ □ Frozen

Önceden eğitilmiş model önce kısa bir eğitmeden geçirilir. Daha sonra asıl eğitime uygulanarak probleme uygun sınıflandırıcı elde edilir.

Önceden Eğitilmiş Modeller

Deneylerde aşağıda sıralanan önceden eğitilmiş modeller kullanılmıştır:

- Inception [10]
- Resnet50 [8]
- VGG16 [6]
- VGGFace [5]

VGGFace VGG Face[5] verikümesi, diğer modeller Imagenet[7] verikümesi ile eğitilmiştir.

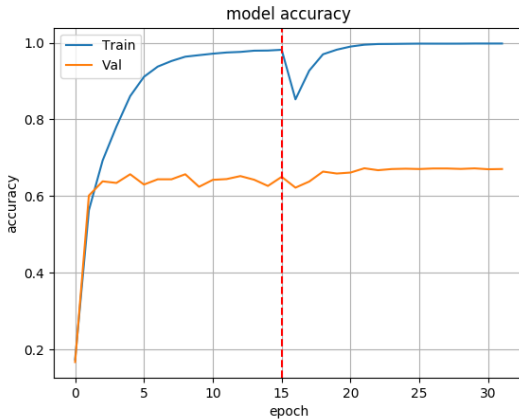
VGGFace for Emotion Recognition

VGGFace modelinin Tam Bağlı katmanları yenileriyle değiştirildi. Softmaks katması 7 adet nöron içerecek biçimde yeniden eklendi.
En iyi sonucu ürettiği için VGGFace modelinden üretilen model baz model olarak kabul edildi.

VGGFace for Emotion Classification
Input(224X224X3)
Conv 64 (3X3)
Conv 64 (3X3)
Max Pooling (2X2)
Conv 128 (3X3)
Conv 128 (3X3)
Max Pooling (2X2)
Conv 256 (3X3)
Conv 256 (3X3)
Conv 256 (3X3)
Max Pooling (2X2)
Conv 512 (3X3)
Conv 512 (3X3)
Conv 512 (3X3)
Max Pooling (2X2)
Conv 512 (3X3)
Conv 512 (3X3)
Conv 512 (3X3)
Max Pooling (2X2)
FC6-1024
FC7-1024
FC8-7

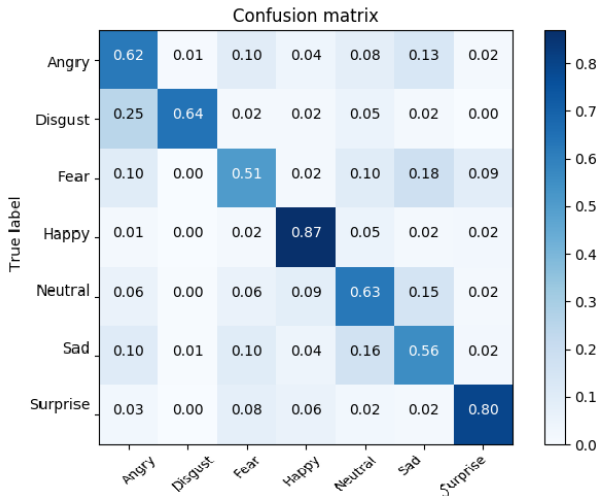
Eğitim

Eğitim&Geçerleme İisabet Eğrisi



Kırmızı çizgi kısa ve uzun yeniden eğitim işlemlerinin sınırını belirtir. VGGFace modelinin eğitimine aittir.

Karışıklık Dizeyi



VGGFace modeline ait karışıklık düzeyi.

Deney Sonuçları

Yeniden Eğitim FER13 Geçerleme Sonuçları

Model	Doğruluk	Kesinlik	Duyarlılık
Inception	0.5668	0.7248	0.7223
ResNet50	0.6120	0.7654	0.7533
VGG16	0.6464	0.8138	0.7586
VGGFace	0.6779	0.8089	0.8072

Karmaşıklığı Azaltma

VGGFace for Emotion Recognition

Son "Max Pooling"
katmanından sonra "Global
Average Pooling" katmanı
eklendi.

"Global Average Pooling"
katmanı tek tek bütün
kanalların ortalamasını alır.
NXNXC boyutlu bir vektör,
C boyutlu hale gelir.

VGGFace for Emotion Classification
Input(224X224X3)
Conv 64 (3X3)
Conv 64 (3X3)
Max Pooling (2X2)
Conv 128 (3X3)
Conv 128 (3X3)
Max Pooling (2X2)
Conv 256 (3X3)
Conv 256 (3X3)
Conv 256 (3X3)
Max Pooling (2X2)
Conv 512 (3X3)
Conv 512 (3X3)
Conv 512 (3X3)
Max Pooling (2X2)
Conv 512 (3X3)
Conv 512 (3X3)
Conv 512 (3X3)
Max Pooling (2X2)
Global Average Pooling
FC6-1024
FC7-1024
FC8-7

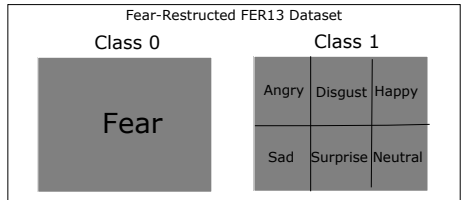
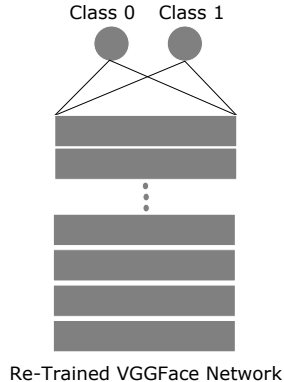
Yeniden Eğitim FER13 Geçerleme Sonuçları

Model	Doğruluk	Kesinlik	Duyarlılık
Inception	0.5668	0.7248	0.7223
ResNet50	0.6120	0.7654	0.7533
VGG16	0.6464	0.8138	0.7586
VGGFace	0.6779	0.8089	0.8072
Basitleştirilmiş VGGFace	0.6949	0.8253	0.8148

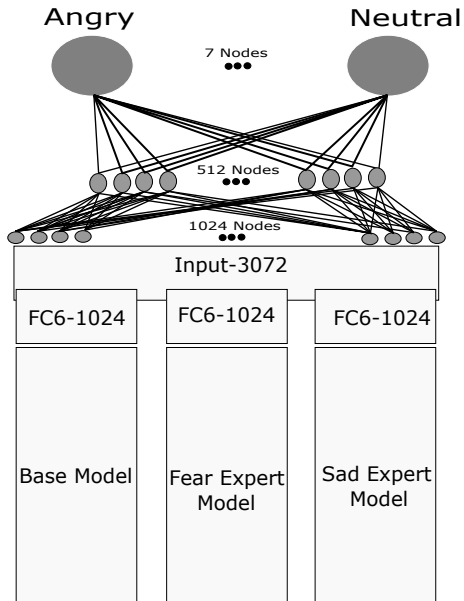
Topluluk Öğrenmesi

Uzman Modellerin Eğitilmesi

Veri Kümesi 1vsAll şeklinde yeniden yapılandırılıp, veri kümesiyle 1 sınıfta uzman sınıflandırıcılar eğitilir. Bu şekilde her sınıf için uzman sınıflandırıcı eğitilir.



Baz model ve seçilmiş uzman modeller kullanılarak 2 gizli katmanlı MLP eğitilir.



Topluluk Öğrenmesi Sonuçları

Topluluk Öğrenmesi En İyi Geçerleme Sonuçları

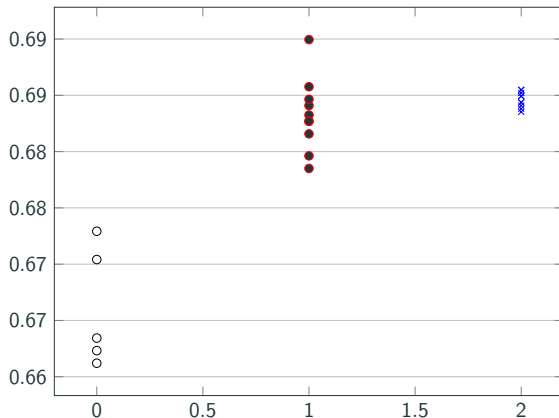
Model	Doğruluk	Kesinlik	Duyarlılık
VGGFace	0.6779	0.8089	0.8072
Basitleştirilmiş VGGFace	0.6949	0.8253	0.8148
Yığın Baz Modeller	0.6662	0.8061	0.7932
Yığın Uzmanlar	0.6905	0.8206	0.8133

Yığın Baz Modeller: Basitleştirilmiş VGGFace + VGG16 + ResNet50

Yığın Uzmanlar: Basitleştirilmiş VGGFace + Korku Uzman + Üzgün Uzman

Topluluk Öğrenmesi Geçerleme Sonuçları




Çember: VGGFace, Nokta: Basitleştirilmiş VGGFace, x: Yığın Uzmanlar











Sonuç ve Değerlendirme




- Transfer Öğrenme önceden eğitilmiş modelleri kullanan etkili bir öğrenme yöntemidir.
- Transfer Öğrenme için probleme uygun olan modelin seçilmesi önemlidir.
- Model karmaşıklığının eldeki veriye göre uygun olması gerekir.
- Farklı kabiliyetlere sahip CNN sınıflandırıcıları Topluluk Öğrenmesi yöntemiyle daha güvenilir sınıflandırıcıların oluşturulması için kullanılabilir.

Sorular?

-  H.Ng, V.D. Nguyen, V. Vonikakis, S. Winkler, “Deep learning for emotion recognition on small datasets using transfer learning”, Proceedings of the 2015 ACM on international conference on multimodal interaction 443–449, 2015.
-  Z. Yu, C. Zhang, “Image based static facial expression recognition with multiple deep network learning”, Proceedings of the 2015 ACM on International Conference on Multimodal Interaction 435–442, 2015.
-  B. Kim, J. Roh, S. Dong, S. Lee, “Hierarchical committee of deep convolutional neural networks for robust facial expression recognition”, Journal on Multimodal User Interfaces Vol 10 pages 173–189, 2016.

-  B. Kim, H. Lee, J. Roh, S. Lee, “Hierarchical committee of deep cnns with exponentially-weighted decision fusion for static facial expression recognition”, Proceedings of the 2015 ACM on International Conference on Multimodal Interaction 427–434, 2015.
-  O.M. Parkhi, A. Vedaldi, Andrea, A. Zisserman, “Deep Face Recognition”, PBMVC Vol 1 page 6, 2015.
-  K. Simonyan, A. Zisserman, “Very deep convolutional networks for large-scale image recognition”, arXiv preprint arXiv:1409.1556, 2014.
-  J. Deng, W. Dong, R. Socher, L. Li, K. Li, L. Fei-Fei, “Imagenet: A large-scale hierarchical image database”, Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on pages 248–255, 2009.

-  K. He, X. Zhang, S. Ren, J. Sun, “Deep residual learning for image recognition”, Proceedings of the IEEE conference on computer vision and pattern recognition pages 770–778, 2016.
-  P.L. Carrier, A. Courville, “FER-2013 face database”, Technical report, 1365, Université de Montréal, 2013.
-  S. Christian , L. Wei, J. Yangqing, S. Pierre, R. Scott, A. Dragomir, E. Dumitru, V. Vincent, R. Andrew, “Going deeper with convolutions”, CoRR Vol. abs/1409.4842, 2015.
-  N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting”, The Journal of Machine Learning Research Vol. 15 pages 1929–1958, 2014.

-  E. Barsoum, C. Zhang, C.C. Ferrer, Z. Zhang, “Training deep networks for facial expression recognition with crowd-sourced label distribution”, Proceedings of the 18th ACM International Conference on Multimodal Interaction pages 279–283, 2016.
-  I. J. Goodfellow, D. Erhan, P.L. Carrier, A. Courville, “Challenges in representation learning: A report on three machine learning contests”, International Conference on Neural Information Processing pages 117–124, 2013.
-  A. Savoiu, J. Wong, “Recognizing Facial Expressions Using Deep Learning” , , 2017.