

GÖRÜNTÜLER ÜZERİNDEN KALABALIK ANALİZİ

Boğaz'da Yapay Öğrenme İsmail Arı Yaz Okulu 2018



Merve Ayyüce KIZRAK

Danışman: Dr. Öğr. Üyesi Bülent BOLAT

Elektronik ve Haberleşme Mühendisliği, Haberleşme Doktora Programı,
YILDIZ TEKNİK ÜNİVERSİTESİ

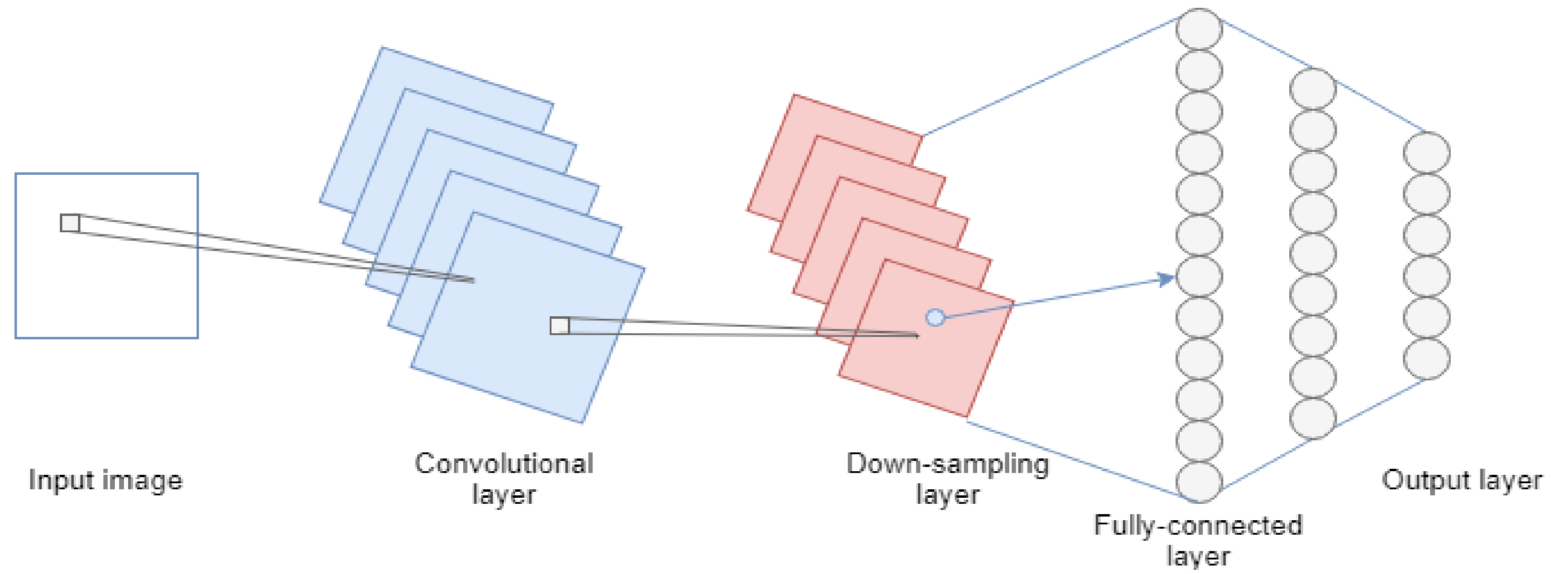
ayyucekizrak@gmail.com

1. MOTİVASYON

Büyük şehirlerde nüfus yoğunluğunun artması şehir güvenliğinin sağlanmasını zorlaştıran önemli bir problem sahasıdır. Sosyal yaşamın modellenmesinde kullanılan, aranmakta olan belirli kişilerin bulunması, kalabalık yoğunluğunun analizi, kalabalık benzetimi, anormal ve tehlikeli durumların saptanıp değerlendirilmesi, işaretlenen kişi ve/veya grupların takibi, kamuya açık alanların tasarlanması, spor müsabakalarının değerlendirilmesi, müze ve kültür-sanat toplantı alanlarının tasarlanması ve giriş-çıkış yönetimi gibi konularda statik görüntü ve video görüntülerinin analiz edilmesi gibi problemlerin gerçek zamanlı olarak çözülebilmesi büyük önem arz etmektedir. Ayrıca arenalar, stadyumlar, sinema, tiyatro salonları ve alışveriş merkezleri gibi insan akışının yoğun olduğu yerlerde de durağan görüntü ve video görüntüleri kullanılarak istatistiksel ve veri analizine yönelik çalışmalar yapılmasına her zaman ihtiyaç duyulmaktadır.

2. YÖNTEM

Kalabalık davranış analizi ve kişi sayımı için derin öğrenme alanındaki son gelişmeler araştırmacılara geleneksel yöntemlerde karşılaşılan birçok zorluğun aşılmasını sağlamaktadır. Özellikle probleme uygun özneteliklerin seçilmesi önemli miktarda deney ve tecrübe gerektirmektedir. Bilgisayarlı görü problemlerinde de kullanılan evrişimli sinir ağları insan görme sistemini en iyi modelleyen özelleşmiş bir yapay sinir ağıdır. Bu yapının daha başarılı sonuçlar vermesi hem öznetelikleri seçmek için bir işlem yapmak gerekmemesi hem de daha derin katmanlarda görüntülere ait üst düzey özelliklerin öğrenilmesinden kaynaklanmaktadır. Böylece araştırmacılar kalabalık sahnelerin analizi için farklı evrişim ve ortaklama (pooling) katmanı konfigürasyonları ile modeller tasarlayarak derin öğrenmeden faydalanmaktadır.



$$ac_{i,j}^l = f \left(\sum_{m=0}^M \sum_{n=0}^N w_{m,n}^l x_{i+m,j+n}^{l-1} + b \right) = f(\text{conv}(w_{m,n}^l, x_{i+m,j+n}^{l-1}) + b^l)$$

Şekil 1: Genel bir evrişimli sinir ağı gösterimi

4. SONUÇLAR

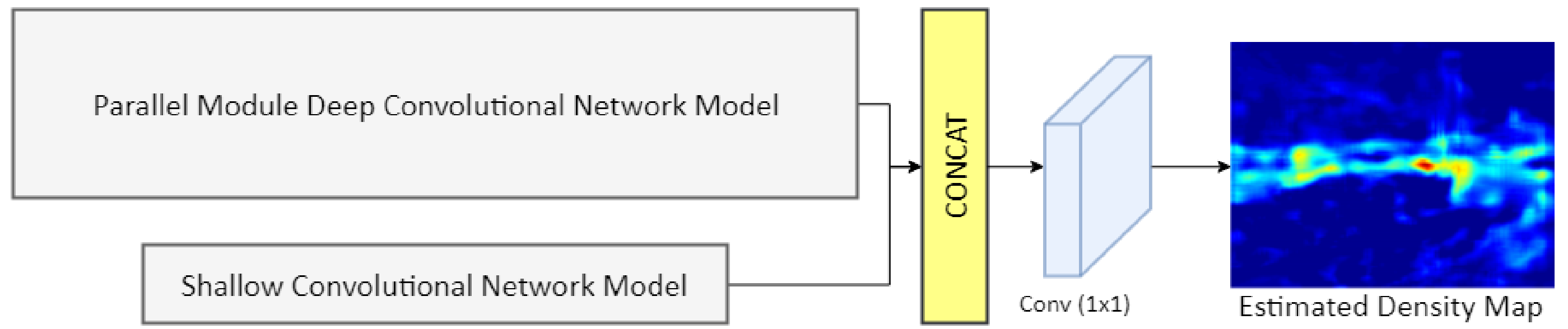
Veri Seti	UCF CC 50 / Shanghai Tech-A
Evrişimli Sinir Ağı Temelli Derin Öğrenme Modeli	MSE: 84.0, MAE: 92.4



Tahmin Edilen Kişi Sayısı=1439, Ground Truth=1428



Tahmin Edilen Kişi Sayısı=646, Ground Truth=634

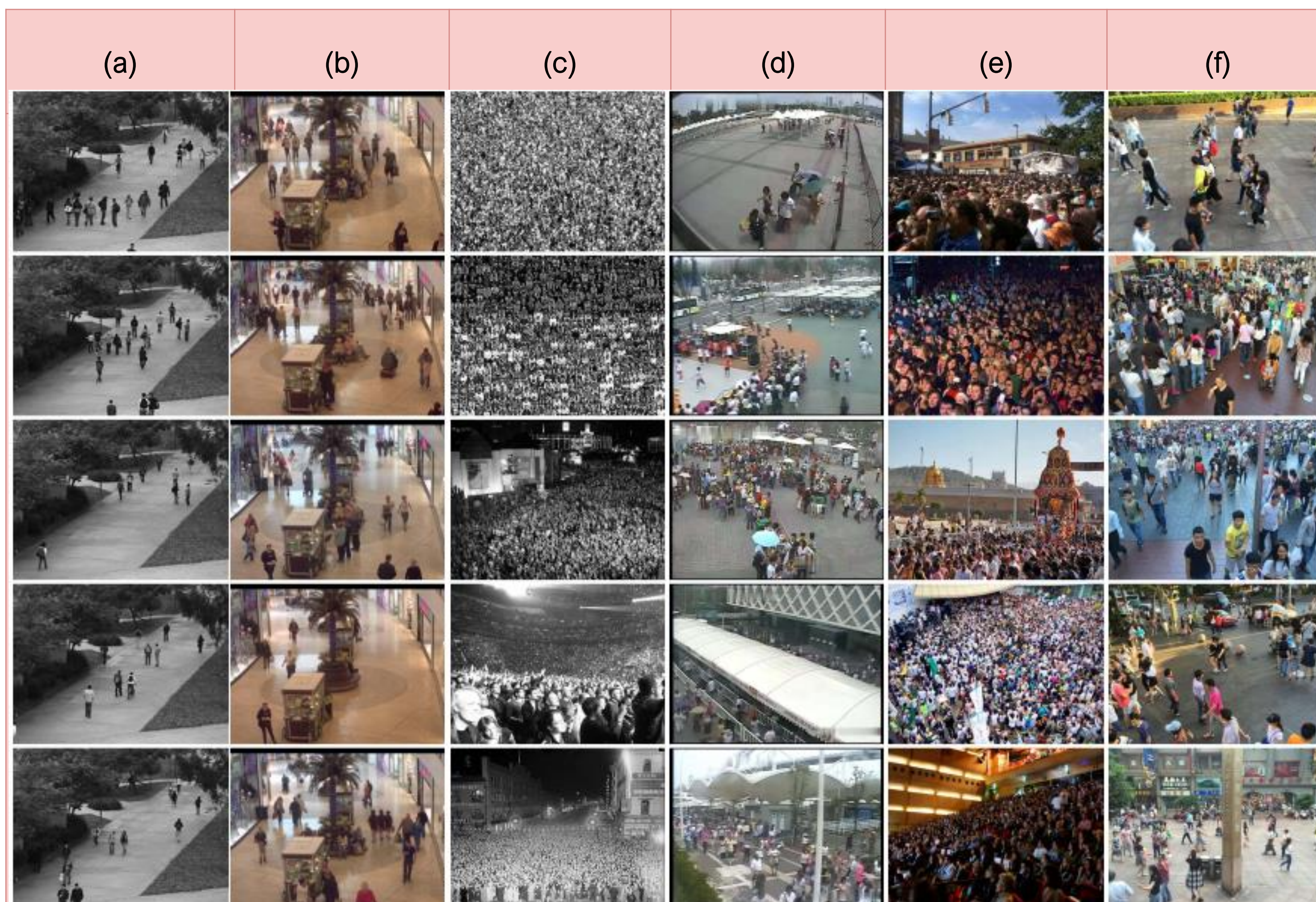


Şekil 2: Tasarlanan iki CNN Modelinin Birleştirilmesi ile Elde Edilen Kalabalık Yoğunluk Haritası

3. VERİ KÜMESİ

Tablo 1: Kalabalık veri kümelerinden örnekler:

UCSD (a), Mall (b), UCF CC 50 (c), WorlExpo'10 (d), ShanghaiTech Part-A (e), ShanghaiTech Part-B (f)



5. DESTEK

Bu doktora tezi NVIDIA Grant Program tarafından desteklenmektedir.

