



**Hacettepe Üniversitesi Bilgisayar Mühendisliği
Bölümü Staj Raporu**

Ad Soyad : Berk Ecer
Öğrenci No : 21228271
Sınıf : 4
Tarih : 01.08.2016 – 16.07.2016
Staj Süresi : 6 hafta/30 iş günü
Stajın Konusu : Görüntü İşleme Veri Kümesi
Staj Yeri : Hacettepe Üniversitesi

Bilgisayarlı Görü Laboratuvarı, Çankaya/Ankara/ Türkiye 06800

İÇİNDEKİLER

1.GİRİŞ	3
2. STAJ YERİ HAKKINDA BİLGİ.....	3
2.1 Amaçlar	3
2.2 Araştırma Alanları	3
2.3 Yürütülen Araştırmalar	4
3. STAJ SÜRESİNCE YAPILAN ÇALIŞMALAR.....	5
3.1 Araştırma ve Ön Çalışmalar	5
3.2 Proje	7
4. SONUÇ VE DEĞERLENDİRME	11
5. KAYNAKLAR	12

1.GİRİŞ

3. sınıf zorunlu yaz stajımı Hacettepe Üniversitesi Bilgisayar Mühendisliği bölümünde bulunan Bilgisayarlı Görü Laboratuvarında yaptım. Stajın ilk haftasında üzerinde uğraşacağım veri kümeleri hakkında bilgi topladım ve özelliklerini araştırdım. Rutin olarak yapılan toplantılara katıldım. Geri kalan haftalarda ise görüntü işlemede kullanılacak bir media(video) veri kümesi oluşturdum.

2. STAJ YERİ HAKKINDA BİLGİ

Hacettepe Üniversitesi bünyesinde yer alan Bilgisayarlı Görü Laboratuvarı (HUVCL)'nin temel hedefi verimli teknikler ve araçlar üreterek makinelerin görsel verileri daha iyi yorumlayabilmesine ve anlamlandırabilmesini mümkün kılmaktır.

2.1. Amaçlar

Hacettepe Üniversitesi Bilgisayarlı Görü Laboratuvarı (HUCVL) ağırlıklı olarak, görsel algılamanın bilgisayar tarafından anlaşılması noktasında araştırma çalışmaları yapmaya odaklanmış, Hacettepe Üniversitesi Bilgisayar Bilimleri ve Mühendisliği bünyesinde kurulu bir araştırma laboratuvarıdır. İlginç problemleri analiz etmek ve çözmek için, Makina Öğrenmesi, Görüntü İşleme ve Bilgisayarlı Görü gibi Bilgisayar Bilimleri'nin temel elementlerini birbiriyle harmanlar.

2.2. Araştırma Alanları

HUCVL'nin araştırma alanları arasında;

- İnsan hareketi tanıma
- Resim ve video işleme
- Büyük ölçekli çokluortam bilgisi çıkarımı

- Düşük seviye görü
- Medikal görüntü işleme ve sınıflama
- Çokluortam verisi madenciliği
- Obje tanıma
- Görsel çekicilik çıkarımı
- Görsel takip gibi konular yer alır.

2.3. Yürütülen Araştırmalar

HUCVL bünyesinde yürütülen bazı araştırma başlıkları şöyledir;

- The Use of Multiple Cues and Contextual Knowledge in Computer
- Learning-based Analysis and Recognition of Human Interactions
- Understanding Images and Visualizing Text: Semantic Inference and Retrieval by Integrating Computer Vision and Natural Language Processing
- Towards A Unified Framework for Finding What Is Interesting In Videos
- City-Wide Video Surveillance System

3. STAJ SÜRESİNCE YAPILAN ÇALIŞMALAR

Staj Yapılan Tarih : 01.08.2016 – 16.09.2016

Mesai Günleri : Pazartesi – Cuma

Stajımın ilk günlerini, Dr. Erkut Erdem ve Dr. Aykut Erdem tarafından verilen, Yahoo Flickr Creative Commons 100M Dataset(YFCC100M)'ini araştırmak ve içeriği hakkında bilgi edinmek ile geçirdim. Bu veri kümesi kullanılarak yapılan projeler ve araştırmalar hakkında makaleler okudum. Edindiğim bilgileri bir sunum haline getirdim. Daha sonra bu veri kümesine benzeyen fakat konu, teme ve lokasyonları belirlenmiş görüntüleri araştırarak yeni bir veri kümesi oluşturmakla uğraştım.

3.1 Araştırma ve Ön Çalışmalar

Yahoo Flickr Creative Commons 100 Million Dataset (YFCC100M)

- YFCC100M hiç yayımlanan en büyük kamu multimedya koleksiyonudur.
- Yaklaşık 99.200.000 resim ve 0,8 milyon video olmak üzere 100 milyon medya nesnesi içerir.
- Sadece meta verileri içeren 12.5GB sıkıştırılmış arşiv olarak Amazon Web Services aracılığıyla dağıtılmaktadır.
- Sürekli gelişmektedir.

Metadata

Metadata, kısaca, bir veri hakkındaki verilerdir. Detaylı olarak açıklanırsa, belirli bir veri setine ya da kaynak hakkında nasıl, ne zaman ve kim tarafından oluşturulduğu hakkında tanımlayıcı bilgiler içerir.

Her nesne:

- Nesnenin Flickr tanımlayıcı tarafından temsili
- Nesneyi oluşturan kullanıcı
- Kullanılan kamera
- Çekildikten ve yüklendiği zaman
- Çekildiği konum (varsa)
- Başlık, açıklama ve etiketler

Etiketler

68552616 fotoğraf ve veri kümesi, kullanıcıları 418507 video etiketleri, ya da anahtar kelime ile açıklamaları vardır. Etiketler, kişi (bebek, aile), kuruluşlar, hayvanlar (kedi, köpek), yerleri (parkı, plaj), seyahat (doğa, şehir) ilgili bilgiler tutarak veri kümesini daha zengin ve kullanışlı hale getirmektedir.

3.343.487 fotoğraf ve 7281 videoları toplam makine etiketleri-etiketler otomatik olarak oluşturulur ve kamera, bilgisayar, uygulama ya da diğer otomatik sistem tarafından eklenen etiketleri de taşırlar.

Zaman Aralığı

- Geçmiş veriler güvenilir değildir.
- Bazı kamera saatleri dahi doğru ayarlanmamıştır.

Bu sebeplerden ötürü güvenilebilir ve kullanışlı bir zaman bilgisi sunmamaktadır.

Konumlar

Veri kümesi içerisinde konumları kullanıcı tarafından veya otomatik olarak GPS yoluyla eklenmiş 48366323 fotoğraf ve 103.506 videolar vardır. Genel olarak, veri kümesi (bu ülkelerde ve adalar gibi) 249 farklı bölgeleri kapsayan ve uluslararası sularda ve uluslararası hava sahasında çekilen fotoğraf ve videoları içerir.

Güçlü Yanları

- + Tasarım: Gerçek hayatı temsilen oluşturulan bir veri kümesidir.
- + Veri eşitliği: Toplanan verilerin kaliteleri ve bazı standartları dengidir.
- + Modellemeleri: Bir çok farklı alanda kullanılabilen modellemeleri vardır.
- + Metadata: Veri özetleri gelişmiştir.

Kısıtlamaları

- Dipnotları(Annotations): Tüm verilerde aynı detay seviyesinde dipnot bulunmamaktadır.

YFCC100M Kullanarak Yapılan Bazı Uygulamalar Aşağıda

Gösterilmiştir:

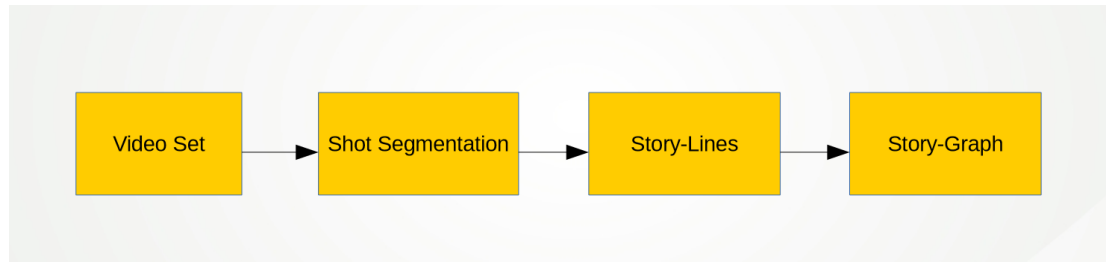
- MediaEval Placing Task: automatically estimate the geo-coordinates
- Towards Total Scene Understanding: Classification, Annotation and Segmentation in an Automatic Framework
- Towards Automatic Extraction of Event and Place Semantics from Flickr Tags

3. 2 PROJE

Bu proje kapsamında daha sonra görüntü işlemede projesinde kullanılacak bir media veritabanı oluşturdum. Bu görüntü işleme görsel özetleme olarak açıklanabilir.

Görsel özetleme çalışması için öncelikle elimizde bir çok farklı görsel verinin bulunması gerekiyor. Daha sonra programın bu karışık veri havuzundan anlamlı özetler çıkarması planlanıyor. Bu işlemler için öncelikle elimizde bir veri kümesinin bulunması gerekiyordu. Ben projenin başlangıcında araştırdığım veri kümelerinin (YFCC100M) üzerine ek olarak belirli konseptlerde video'lar bularak yeni bir veri kümesi oluşturdum.

Görsel özetleme projesinin adımlar temel olarak şu şekilde gösterilebilir.



Projenin ilk adımı olarak bir video kümesine ihtiyaç duyuluyordu. Bu noktada beraber çalıştığım araştırma görevlisi ve öğretmenlerimizle bu video kümesi hakkında konuşup video araştırma ve toplama işine başladım.

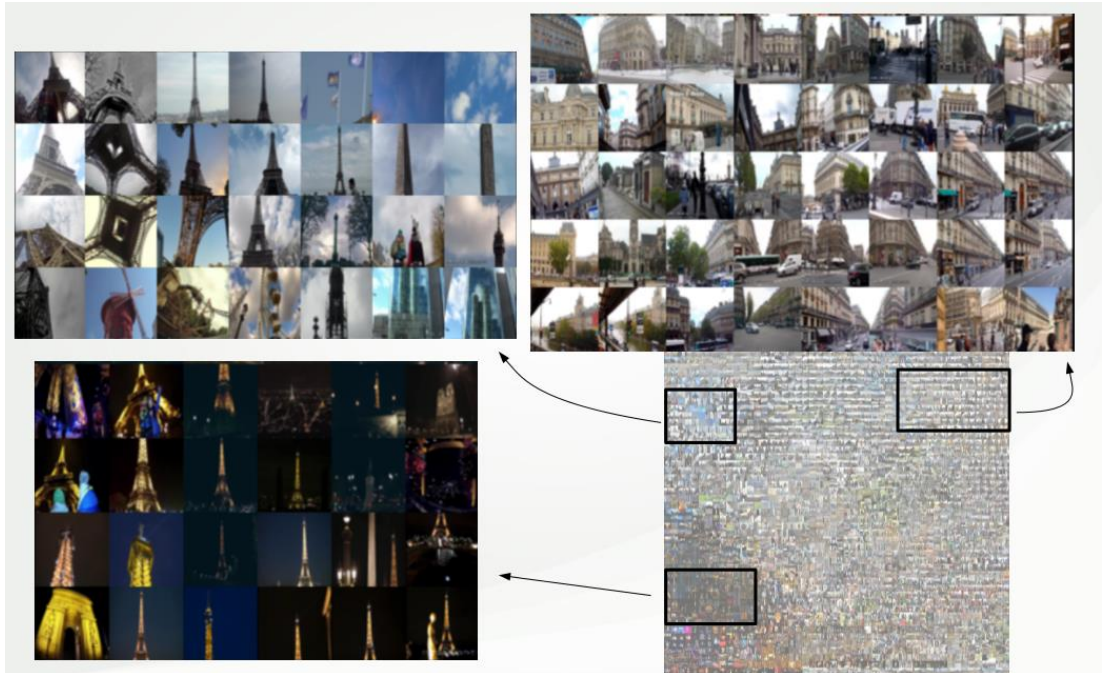
Gerekli videoların bazı kıstasları bulunmaktaydı. Öncelikle bu verilerin konumları belirlendi. Örnek olarak New York, Londra ve Roma 'da çekilmiş videolar araştırıldı.

Bu videoların konumlarının doğru ve güvenilir olması gerekiyordu. Bu sebepten ötürü Geo Search Tool kullanılarak gerekli videolar bulundu. Geo Search Tool, konum ve zaman aralığı girdisiyle Youtube üzerinden video sorgulama yarayan bir programdı.

Bu sayede videoların konumları hakkında güvenilir bir bilgiye ulaşmış olduk. Doğrudan Youtube üzerinden video aratmamızın sebebi ise konum bilgilerinin güvenilirliğinin zayıf olmasıydı.

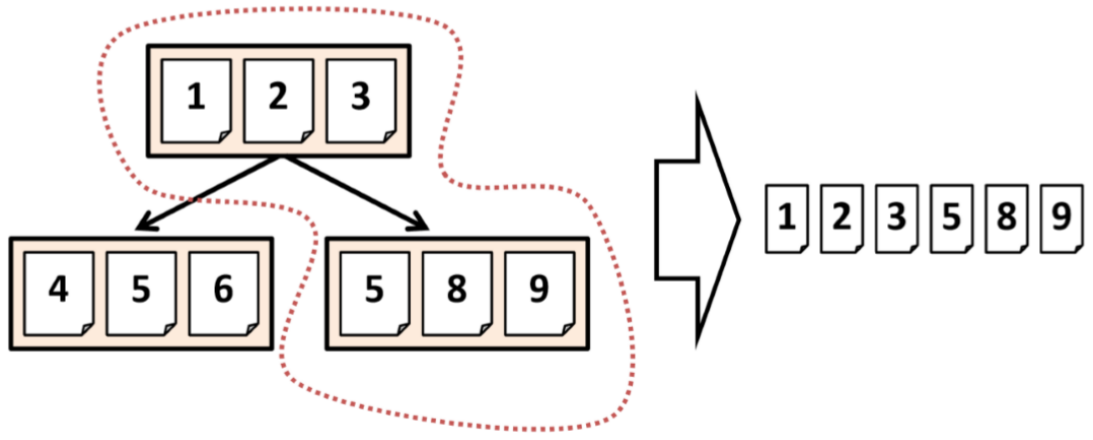
İkinci kıstas olarak videoların içerikleri de bizim için önem arz ediyordu. Videolarda herhangi bir altyazı veya bilgilendirici açıklamanın olmaması gerekiyordu. Bunun sebebi ise videolar kullanılarak shotlara ayrılacaktı ve bu yazıların görüntü işleme aşamasını zorlaştırdığı söylenmişti.

Videoların uzunlukları ve görüntü kalitesi de bizim için önemliydi. Videolar çok kısa olursa istenildiği kadar shot çıkmayabilirdi bu yüzden ortalama olarak 15 dakika ve üzeri videolar indirildi. Görüntü kalitesi düşük videolar ise görüntü işlemeyi zorlaştıracığından 480p ve üzeri videolar tercih edildi.



Şehir başına ortalama 3 saatlik video havuzu indirildikten sonra. Videolar izlenip istenilen şartlara uygunluğu gözlemlendi.

Daha sonra videoları shotlarına ayıracak kod araştırıldı ve araştırma görevlisinin yazmış olduğu kod incelendi. Staj süresinin bitmesi sebebiyle shotlarına ayırma işlemi tamamlanmadı.



Videolar shot'larına ayrıldıktan sonra bu resimleri anlamlı bir hikaye oluşturacak şekilde birleştirme ve bir ağaç yapısı oluşturulacaktı.

4. SONUÇ VE DEĞERLENDİRME

Yaptığım staj boyunca görüntü işleme teknikleri ve bir görüntü işleminde kullanılacak veritabanının ne gibi özelliklere sahip olması gerektiğini öğrendim.

Staj öncesi yaptığım lineer cebir ve olasılık çalışmalarıyla eksik bilgilerimi tamamladım. Öncelikle lineer cebirin görüntü işlemedeki önemini anladım. Bu noktada her bir resim veya görüntünün bir matris olarak değerlendirilebileceğini ve temel cebir operasyonlarının ne kadar önemli olduğunu anladım.

YFCC100M medya veritabanının sunduğu olanakları araştırdım ve ne gibi projelerde kullanılmaya uygun olduğunu öğrendim. Bu veritabanının avantajını ve dezavantajlarını tespit ettim.

Görsel özetleme projesi hakkında bir çok farklı bilgi öğrendim ve bu temadaki bir projenin nasıl aşamalarla yapılabileceğini öğrendim.

Görüntü işleme projeleri için veritabanının ne kadar önem arz ettiğini anladım ve en başından doğru belirlenmiş adımlara uyarak daha sonra problem yaşamamak gerektiğini anladım. Veritabanı hazırlarken kriterlerin düzgün belirlenmesi gerektiğini ve ileride karşımıza problem çıkarması için yapılması gereken şeyleri öğrendim.

Görüntü işleme laboratuvarında bir araştırma takımının nasıl olduğunu görev dağılımını ve haftalık sunumlar insanlarla nasıl birlikte çalışılabileceğini anladım ve araştırma ortamının içinde bulunma imkanına sahip oldum.

5. KAYNAKLAR

- <http://vision.cs.hacettepe.edu.tr>
- http://delivery.acm.org/10.1145/2670000/2661125/p27-choi.pdf?ip=193.140.236.193&id=2661125&acc=CHORUS&key=956257EA1AE17323%2E7707A0CE88B0A386%2E4D4702B0C3E38B35%2E6D218144511F3437&CFID=651864775&CFTOKEN=18873284&__acm__=1470303609_f261457d4e3ac040c712acef3ad3a3ca
- <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5206718>
- http://delivery.acm.org/10.1145/1280000/1277762/p103-rattenbury.pdf?ip=193.140.236.193&id=1277762&acc=ACTIVE%20SERVICE&key=956257EA1AE17323%2E7707A0CE88B0A386%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&CFID=651864775&CFTOKEN=18873284&__acm__=1470308106_9cda06aa5fe1410dffd6fcc6eaa6887c
- <http://www.geosearchtool.com>