DÜMF Mühendislik Dergisi

web: http://dergipark.gov.tr/dumf

Derin öğrenmenin Caffe kullanılarak grafik işleme kartlarında değerlendirilmesi

Ayşegül UÇAR*
Fırat Üniversitesi, Mekatronik Mühendisliği Bölümü, Elazığ
agulucar@firat.edu.tr , Tel: (424) 237 00 00 (6358)

Mehmet Safa BİNGÖL
Ömer Halisdemir Üniversitesi, Mekatronik Mühendisliği Bölümü, Niğde

Geliş: 13.03.2017 , Kabul Tarihi: 20.10.2017

Öz

Yapay Sinir Ağları (YSA'lar) bilgisayar görmesinden konuşma tanımaya kadar birçok uygulamada kullanılmaktadır. YSA'lar, geleneksel hesaplama yöntemlerinden daha üstün sonuçlar vermektedir. Ancak veri sayısının büyük olduğu uygulamalarda, çok fazla hesaplama gerektirmektedir. Bu nedenle verileri paralel işleyerek sistemlere hız kazandıran Grafik İşleme Birimlerinin kısaca GPU'ların (Graphics Processing Units) kullanımını zorunlu hale gelmiştir. Son yıllarda, YSA'ların bir çeşidi olan derin öğrenme algoritmaları GPU'lar sayesinde, gerçek hayattaki birçok uygulamada başarılı olarak uygulanmış ve GPU içeren gömülü sistemlere talep artmıştır.

Bu çalışmada, derin öğrenme yöntemlerinden birisi olan Derin Konvolüsyonel Sinir Ağları (DKSA'lar) ve katmanları kısaca tanıtılmış ve düşük maliyetli Nvidia Jetson TK1/TX1 gömülü sistemleri genel olarak incelenmiştir. Derin öğrenme algoritmalarını uygulamak için Kaliforniya Üniversitesi, Berkeley tarafından geliştirilen Caffe programı kullanılmıştır. Son olarak paralel işleme gücüne sahip olan Nvidia Jetson TK1/TX1 kartları üzerinde MNIST verileri kullanılarak LeNet ağının eğitimi gerçekleştirilmiştir. Eğitim için kartların hem Merkezi İşleme Birimleri kısaca CPU'ları (Central Processing Unit) hem de GPU'ları kullanılmıştır. Deneysel sonuçlar için, Nvidia Jetson TK1/TX1 kartlarına ek olarak Nvidia GTX550 ve GTX960 ekran kartlarına sahip iki bilgisayar da kullanılmıştır. Sonuçlar, hız ve doğruluk açısından değerlendirilmiştir.

Anahtar Kelimeler: *ECoG Derin öğrenme; konvolüsyonel sinir ağları; LeNet; MNIST; Caffe; Nvidia Jetson TK1/TX1*

* Yazışmaların yapılacağı yazar

DOI:

Giriş

Günümüzde veri kümelerinin büyümesi ile Merkezi İşleme Birimleri kısaca CPU'lar (Central Processing Unit) üzerinde yapılan hesaplamalar yetersiz kalarak yerini verileri paralel işleyen Grafik İşleme Birimlerine kısaca GPU'lara (Graphics Processing Units) bırakmıştır. Bu sayede derin öğrenme hızlı bir şekilde gerçek hayattaki uygulamalarda kullanılmaya başlanmıştır. Özellikle, çok sayıda eğitim verilerinden öznitelik çıkarma işlemi yapabilen sistemler oluşturmak için ileri teknolojide, derin ileri beslemeli Yapay Sinir Ağları'nın (YSA'ların) kullanımı artmıştır. Çok katmanlı ileri beslemeli YSA'ların bir türü olan Derin Konvolüsyonel Sinir Ağları (DKSA'lar), LeCun tarafından 1989 yılında bilgisayar görmesi uygulamaları için tanıtılmıştır (LeCun vd., 1989).

DKSA'larda, görüntüler ağa giriş olarak verilir. Herhangi bir öznitelik çıkarma yöntemi kullanmadan, sadece oluşturulan öznitelik hiyerarşisi otomatik olarak öğrenme için kullanılır. Hiyerarşik yaklaşım daha karmaşık yapılı ve çok katmanlı özniteliklerin öğrenilmesine izin verir. Hiyerarşik yaklaşım, yüksek katmanlarda dönme ve bozulmaya duyarsız öznitelikler elde etmeyi ve daha karmaşık genellemeleri elde etmeyi sağlar. Öznitelik haritalarının hiyerarşisi, öğrenilmiş filtre ile giriş görüntüsünü konvüle ederek oluşturulur.

Derin öğrenmede eğitim, geleneksel çok katmanlı algılayıcıdaki ile karşılaştırıldığında zor olarak görünmesine rağmen, geleneksel yöntemler kullanılarak daha kolay eğitilebilir. Bunun nedeni, gerçek model karmaşıklığını azaltan ağırlık paylaşımı ilkesinin DKSA'larda kullanılmasıdır (Bengio, 2009). Bu özellik ile birlikte DKSA'ların mimarisinden dolayı özel girişler için farklı konvolüsyonlar tanımlanır. Bu özellik, günümüzde DKSA'lar ile derin öğrenme konusunu aktif bir araştırma konusu yapmıştır.

DKSA'lar, otonom araçlar, konuşma tanıma, görüntü altyazılama, doğal dil işleme, görüntü analizi, kişi, nesne veya sahne tanıma, büyük ölçekli video analizi ve kanserli hücrelerin tespiti gibi birçok alanda kullanılmaktadır (Uçar vd., 2016; Christian vd., 2013).

Bu çalışmada, DKSA'ların katmanları kısaca anlatıldıktan sonra Nvidia Jetson TK1/TX1 kartları kısaca tanıtılmıştır. Daha sonra Nvidia Jetson TK1/TX1 kartları ve Nvidia GTX550 ve GTX960 ekran kartlarına sahip iki bilgisayar üzerinde LeNet ağının eğitimi gerçekleştirilmiş ve sonuçlar, doğruluk ve zaman bakımından değerlendirilmiştir. GPU'ların CPU'lara göre üstünlükleri ve gömülü sistemlerin getirileri değerlendirilmiştir.

Derin Konvolüsyonel Sinir Ağları

DKSA'larda, her katmanın giriş ve çıkışları görüntü dizi kümeleri ile gösterilir. DKSA'lar, konvolüsyonel katman, nöron öbeklerinin çıkışlarını birleştiren bölgesel veya yerel öznitelik seçme katmanı ve parça parça doğrusal olmayan aktivasyon fonksiyonlarına sahip tamamen bağlı katmanların farklı kombinasyonlarından oluşur. DKSA'lar, Şekil 1'de gösterildiği gibi, bitişik katmanların nöronları arasında yerel bir bağlantı modeli uygulayarak uzaysal olarak yerel ilinti kullanmaktadır. DKSA'ların katmanları aşağıda kısaca açıklanmıştır:

Konvolüsyon Katmanı

Konvolüsyon katmanı, DKSA'ların temel katmanıdır (Deng ve Yu, 2014; Christian vd., 2013). Giriş görüntüsü, öğrenebilir filtreler veya çekirdek kümeleri kullanılarak konvüle edilir. Her biri, çıktı görüntüsünde bir öznitelik haritası üretir. Elde edilen öznitelik haritaları, ikinci konvolüsyon katmanının girişine verilir.

Öznitelik Seçici Katman

Öznitelik seçici katman, öznitelik boyutunu azaltır. Giriş görüntüleri üst üste binmeyen dikdörtgen bölgelere bölünür. Bölünen bölgeler, en büyük veya ortalama gibi doğrusal olmayan

alt örnekleme işlemi ile örneklenir. Öznitelik seçici katman, dönme ve bozulmaya karşı az değişim, daha hızlı yakınsama ve daha iyi bir genelleme sağlar. Öznitelik seçici katmanlar, uzaysal boyutu azaltmak için genellikle ardışıl konvolüsyon katmanları arasına yerleştirilir.

Doğrultulmuş Doğrusal Birim (DDB) Katmanı

Doğrultulmuş Doğrusal Birim (DDB) katmanı, doğrultucu olarak çalışan birimleri içerir. YSA'ların bilimsel yazınında, bir x girişi nörona verildiği zaman, doğrultulmuş çıkış

$$f(x)=\begin{cases} x & x \geq 0 \\ 0 & diger \end{cases} \quad (1)$$

olarak tanımlanır. Bu fonksiyon, karar fonksiyonunun doğrusalsızlığını artırır. Aktif biriminin gradyeni 1, diğerleri sıfırdır. Geriye yayılımda gradyenin kaybolma problemi yoktur. En belirgin öznitelikler kuvvetlendirilir, diğer öznitelikler ise bastırılır. Seyrek modeller elde edilir. Doğrultucu fonksiyona düzgün bir yaklaşım,

$$f(x)=ln(1+e^x) \quad (2)$$

olarak verilir. Bu fonksiyonun türevi lojistik fonksiyondur ve genellikle öğrenmeyi hızlandırmak için kullanılır.

Tam Bağlı Katman

Tam bağlı katman, konvolüsyon katmanları, öznitelik seçici katmanlar ve DDB katmanlardan sonra yerleştirilir. Tam bağlı katmanın nöronları, önceki katmandaki tüm aktivasyonlara tamamen bağlıdır. Tam bağlı katman, sınıflandırıcıyı besleyen son öznitelik seçici katman olarak kabul edilir. Aktivasyonları hesaplamak için matris çarpımı ve bir eşik (bias) ilavesi kullanılır.

Değerlendirme Katmanı

Tahmin edilen ve gerçek sınıf değerleri arasındaki ödünleşim, bu katmanda tanımlanan fonksiyonu ile belirlenir. Değerlendirme katmanı genellikle ağın son katmanıdır. Farklı uygulamalar için farklı değerlendirme fonksiyonları mevcuttur. Öklid değerlendirme fonksiyonu, genellikle -∞ ve ∞ aralığında gerçek değerli çıktılar üreterek, regresyon uygulamaları için kullanılır. Öklid fonksiyonu,

$$l(.)=\frac{1}{2}(s_i-y_i)^2 \quad (3)$$

ile tanımlanır. Burada, y hedef değerler ve s ağ çıkışlardır.

Softmax olarak bilinen yaklaşık en büyük değerlendirme fonksiyonu ise, sınıflama uygulamaları için kullanılır. Softmax fonksiyonu (4)'deki gibi kullanılır

$$p_c=\frac{e^{sc}}{\sum_{i=1}^{C} e^{s_i}}. \quad (4)$$

Burada; p_c c sınıfına ait s vektörünün olasılığını gösterir. Softmax fonksiyonu, (5)'deki çapraz entropi değerlendirme fonksiyonunu tanımlamak için kullanılır

$$l(.)=-\sum_{c=1}^{C} y_c\, logp_c\,. \quad (5)$$

Burada; y gerçek sınıf etiketleridir.

Bir DKSA yapısı ve etiketli veri kümesi $S=\{(x_i,y_i)\}_{i=1}^{m}$ verildiği zaman, öğrenme algoritması (6)'daki eniyileme probleminin çözümünü, değerlendirme fonksiyonunu enazlayarak bulur.

$$min_W \sum_{i=1}^{m} l(f(W;x_i),y_i) \quad (6)$$

Burada; W konvolüsyon ve doğrusal katmanların ağırlıklarını göstermektedir.

Bu eniyileme probleminin bir çözümü, eğim iniş yöntemleri kullanılarak elde edilir. DKSA'larda genellikle birinci dereceden türev kullanan ardışıl yöntemler kullanılır. Değerlendirme fonksiyonunun, gradyen ile belirlenen en dik iniş yönünde, hareket edilerek ağrılıklar güncellenir. (7)'de, t anındaki ağırlık kullanılarak, t+1 anındaki ağırlığın nasıl güncellendiği gösterilmiştir.

$$W^{(t+1)} = W^t - \eta_t \nabla_W \left(\sum_{i=1}^{m} l(f(W; x_i), y_i) \right). \quad (7)$$

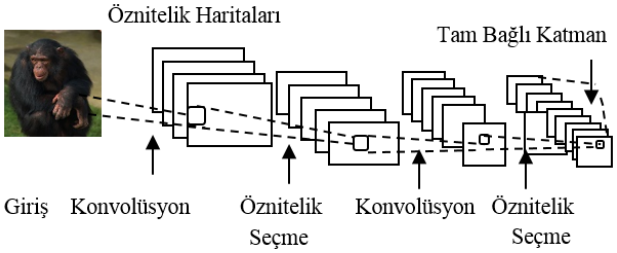
Burada; ∇_W W'ye göre gradyeni ve η_t pozitif skaler öğrenme oranını göstermektedir. DKSA'da η_t parametresi zamana bağlı bir fonksiyon olarak tanımlanabilir.

Gradyenin örneksel olarak hesaplanması, S'in tüm elemanlarının toplanması nedeniyle çok maliyetlidir. DKSA'larda stokastik eğim iniş yöntemi kullanılır. Bu yöntem daha az maliyetlidir (Haykin, 2008). Her bir adımda sabitlenmiş küçük veri kümeleri $I_t \subseteq \{1, \ldots, m\}$ rasgele seçilir ve güncelleme (9)'daki gibi yapılır

$$Z^{(t+1)} = \mu Z^{(t)} - \eta_t \nabla_W \left(\sum_{i \in I_t} l\left(f\left(W^{(t)} + \mu Z^{(t)}; x_i\right), y_i\right) \right) \quad (8)$$

$$W^{(t+1)} = w^{(t)} + Z^{(t+1)}. \quad (9)$$

Burada; $Z^{(t)}$ yardimci vektör ve μ skaler momentum parametresidir.



Şekil 1. Genel bir DKSA mimarisi

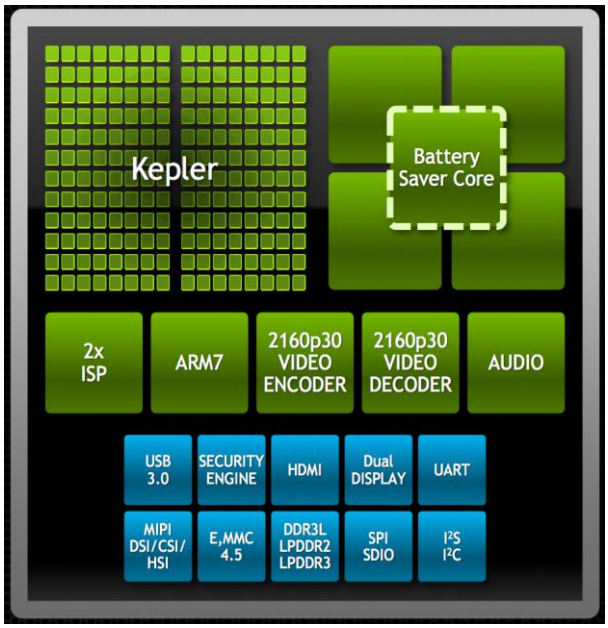
Nvidia Jetson TK1/TX1 Kartları

Jetson TK1/TX1, Nvidia tarafından bilgisayar görmesi, otomotiv, tıp ve robotik uygulamalarında kullanılmak için geliştirilen kartlardır (Nvidia websitesi).

Nvidia Jetson TK1/TX1 geliştirme kartları, tümleşik sistem uygulamalarında, GPU'nun gücünden ve getirdiği kazançlardan yararlanmak için gereksinim duyulan her şeyi sunmaktadır. Kart üzerinde 192 tane CUDA çekirdeği bulunmaktadır. Böylece, CUDA kütüphanesini kullanarak işlemleri paralelleştirmek mümkün hale gelmektedir. Bu sayede gerçek zamanlı görüntü işleme uygulamalarında, zaman açısından büyük bir kazanç elde edilmektedir.

Nvidia Jetson TK1

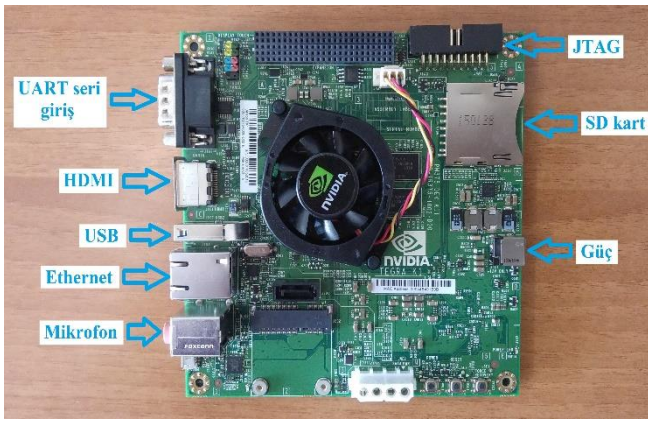
Jetson TK1, 2014 yılının mayıs ayında Nvidia tarafından kullanıma sunulmuştur. Bu kart, hibrit işlemci olan Tegra K1 tabanlıdır. Nvidia Jetson TK1, batarya koruma çekirdeği, çift çekirdekli ARM Cortex A15 CPU ve 192 Nvidia Kepler CUDA çekirdeğine sahiptir. Şekil 2'de Jetson TK1'in mimarisi gösterilmiştir (Bauman, 2015).



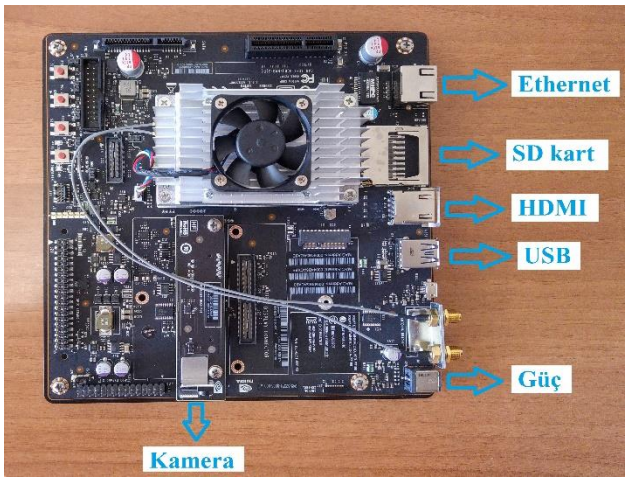
Şekil 2. Nvidia Jetson TK1'in mimarisi

Şekil 3'te Jetson TK1 ve bağlantıları verilmiştir. Kart üzerindeki bağlantılar aşağıdaki gibidir:

- 1 USB 3.0 girişi,
- 1 HDMI girişi,
- 1 mini-PCIe yuva,
- 1 SD/MMC bağlayıcı,
- 1 RS232 seri bağlantı,
- 1 GigE LAN,
- 1 SATA girişi,
- SPI 4 MB boot flash.



Şekil 3. Nvidia Jetson TK1 ve bağlantıları



Şekil 4. Nvidia Jetson TX1 ve bağlantıları

Nvidia Jetson TX1

Jetson TX1, bir modül üzerindeki ilk süper bilgisayardır (Nvidia websitesi). Görsel hesaplama uygulamaları için performans ve güç verimliliği sağlamaktadır.

1 TeraFLOP ve 256 CUDA çekirdeğine sahiptir. 64-bit CPU'lar sayesinde 4K video kodlama ve kod çözme işlemleri yapılabilmektedir. Üzerindeki 1400 MPix/s kapasiteli kamera ve yüksek performanslı GPU'lar sayesinde görüntü işleme içeren derin öğrenme uygulamalarını gömülü olarak yapmaya izin verir.

Şekil 4'te Jetson TX1 ve bağlantıları gösterilmiştir. Kart üzerindeki bağlantılar aşağıdaki gibidir:

- USB 3.0,
- USB 2.0 Micro AB (kurtarma ve ana bilgisayar modunu destekler),
- HDMI,
- M.2 Key E,
- PCI-E x4,
- Gigabit Ethernet,
- Tam Boyut SD,
- SATA Verileri ve Güç,
- GPIOs, I2C, I2S, SPI,
- TTL UART ve Akış Kontrolü,
- Ekran Genişletme Başlığı,
- Kamera Genişletme Başlığı.

Caffe

Caffe, derin öğrenme algoritmalarını uygulamak için Kaliforniya Üniversitesi, Berkeley tarafından geliştirilen açık kaynaklı bir platformdur. Kodlar, GPU hesaplama için kullanılan CUDA kütüphanesi ile C++'da yazılır. Ayrıca CUDA kütüphanesi, Python/Numpy ve MATLAB'ı destekler. Hızlı CUDA kodu ve GPU'nun hesaplama gücü sayesinde, K40 veya Titan GPU kullanılarak her gün 40 milyonun üzerinde görüntü işlenebilmektedir. Bu da yaklaşık olarak her görüntü için 2.5 ms zaman ayrıldığı anlamına gelmektedir (Jia vd., 2014).

Caffe, ilk olarak görüntü işleme için tasarlanmasına rağmen daha sonradan ses tanıma, robotik, astronomi ve sinir bilimi için kullanılmıştır. Caffe'nin, bu hızlı yayılımını sürdürmesiyle farklı bilim alanlarında ve endüstride yaygın olarak kullanılarak derin öğrenmenin getirilerinden faydalanılması beklenmektedir.

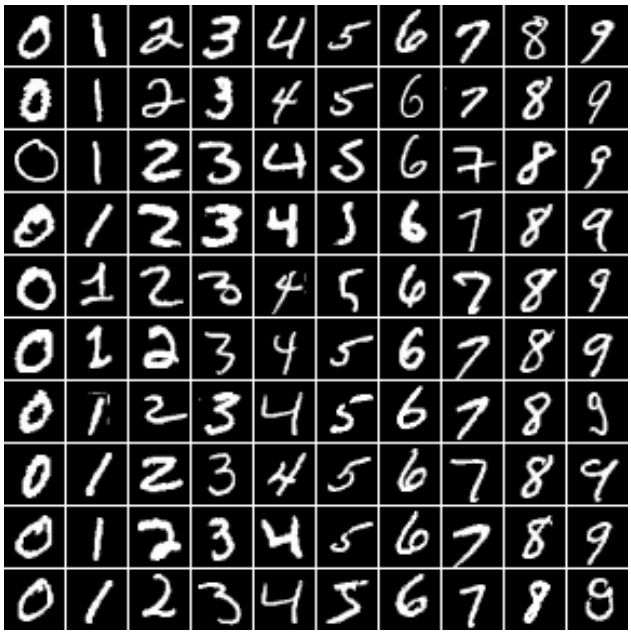
Deneysel Sonuçlar

Bu çalışmada, Intel i7, 3.4 GHz işlemciye sahip, 16 GB Ram ve Nvidia GTX550 ekran kartı olan masaüstü bilgisayar, Intel i7-3.6 GHz işlemciye sahip, 16 GB Ram ve Nvidia GTX960 ekran kartı olan dizüstü bilgisayar ve Nvidia Jetson TK1/TX1 geliştirme kartları kullanılmıştır.

Uygulama olarak el yazısı tanıma yapılmıştır. MNIST el yazısı veri kümesi ve LeNet ağı kullanılmıştır.

MNIST Veri Kümesi

MNIST veri kümesi el yazısı rakamlardan oluşan 28x28 boyutunda 60000 eğitim görüntüsü ve 10000 test görüntüsü içermektedir (LeCun ve Cortes, 1998). 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 el yazısı rakamlarından oluşan MNIST veri kümesinden örnek bir görüntü Şekil 5'te gösterilmiştir.



Şekil 5. MNIST veri kümesinden bir örnek
(LeCun ve Cortes, 1998)

LeNet Ağı

LeNet, konvolüsyonel ağların başarılı olarak uygulandığı ilk ağ yapısıdır. LeNet, 1990 yılında Yann LeCun tarafından geliştirilmiştir (LeCun vd., 1989). LeNet mimarisi başka uygulamalar için kullanılmakla birlikte, daha çok rakamlar için kullanılmaktadır. Şekil 6'da bir LeNet ağının örnek bir gösterimi verilmiştir. Tablo 1'de LeCun tarafından oluşturulan LeNet ağının detaylı mimarisi verilmiştir (LeCun vd., 1989).

Caffe'de Kullanılan Ağ ve Parametreleri

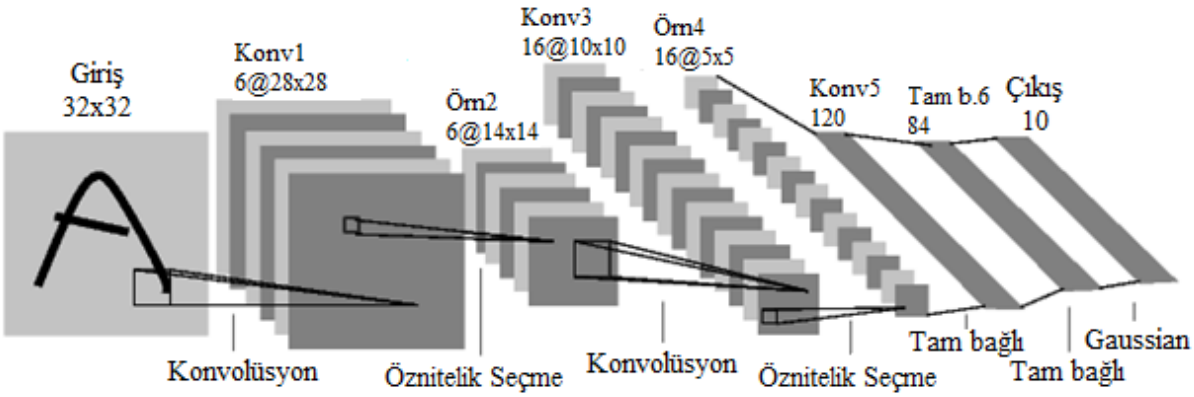
Bu çalışmada, Berkeley tarafından derin öğrenme için geliştirilen Caffe programı kullanılmıştır. Caffe'de, MNIST veri kümesi ile LeNet ağı eğitilirken aşağıdaki parametreler kullanılmıştır. MNIST verileri için özelleştirilmiş ağ yapısı Şekil 6'da, blok diyagramı olarak gösterimi ise Şekil 7'de verilmiştir. MNIST veri kümesinde 10 tane sınıf olduğu için ağ çıkışı da 10 tanedir. Bu çalışmada kullanılan ağ parametrelerinin değerleri Tablo 2'de verilmiştir.

Tablo 1. LeNet ağının mimarisi

Katman	LeNet Modeli	Görüntü Boyutu
0	Giriş (Gri ölçek veya renkli)	3x32x32
1	Konvolüsyon (5x5, 108 harita)	108x28x28
2	Doğrusalsızlık (Doğrultulmuş tanjant)	108x28x28
3	Öznitelik Seçici (Maks.) (2×2, kaydırma 2)	108x14x14
4	Konvolüsyon (5x5, 108 öznitelik haritası)	108x10x10
5	Doğrusalsızlık (Doğrultulmuş tanjant)	108x10x10
6	Öznitelik Seçici (Maks.) (2×2, kaydırma 2)	108x5x5
7	Tam bağlı (100 birim)	100x1x1
8	Doğrusalsızlık (Doğrultulmuş tanjant)	100x1x1
9	Tam bağlı (100 birim)	100x1x1
10	Doğrusalsızlık (Doğrultulmuş tanjant)	100x1x1
11	Tam bağlı (43 birim) + softmax	43x1x1

Tablo 2. Caffe'de kullanılan ağ parametrelerinin değerleri

Parametre	Değeri
Temel öğrenme oranı (Base learning rate)	0.01
Momentum	0.9
Ağırlık gecikmesi (Weight decay)	0.0005
Yığın boyutu (Batch size)	64
Gama (Gamma)	0.0001
Güç (Power)	0.75
En büyük adım sayısı (Maximum iteration)	10000
Test işleminin kaç adımda bir gerçekleştirileceği	500



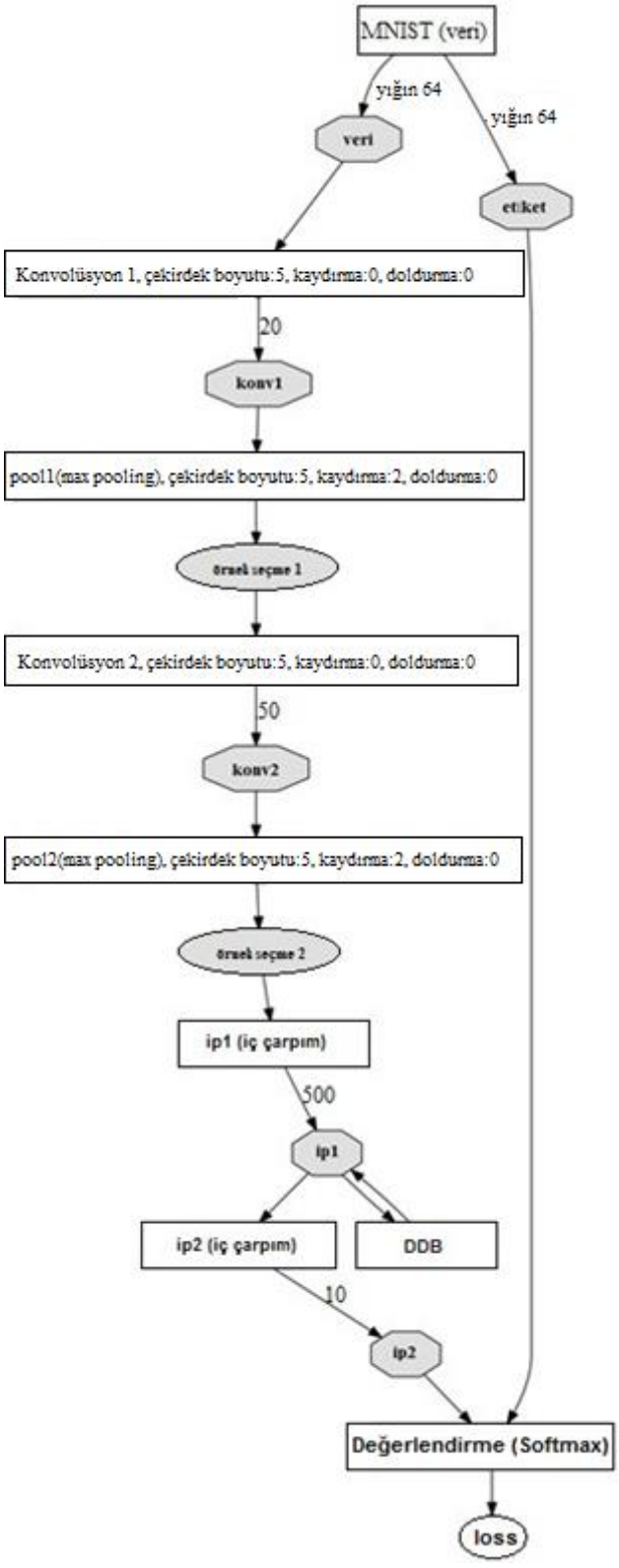
Şekil 6. LeNet ağı (LeCun vd., 1995)

Tablo 3. Deneysel sonuçlar

Cihaz	İşlemci Türü	Süre (dakika)	Doğruluk (%)
Nvidia Jetson TK1	CPU	65	99.20
	GPU	13	99.07
Nvidia Jetson TX1	CPU	24	99.08
	GPU	3	98.98
Nvidia GTX550 Masaüstü Bilgisayar	CPU	14	99.02
	GPU	5	99.08
Nvidia GTX960 Dizüstü Bilgisayar	CPU	12	99.02
	GPU	1	99.14

Tablo 3'de elde edilen sonuçlar gösterilmiştir. Bu sonuçlara göre; ağ eğitimini en hızlı, Nvidia GTX960 içeren dizüstü bilgisayarın GPU işlemcisi yapmıştır. En yavaş ağ eğitimini ise, Nvidia Jetson TK1'in CPU işlemcisi yapmıştır. Diğer taraftan, en yüksek doğruluk oranı Jetson TK1 ile elde edilmiştir.

Nvidia Jetson TX1 ve TK1 birbiriyle karşılaştırtıldığında, GPU ve CPU kullanıldığı zaman, TX1'in TK1'den sırasıyla 4.33 ve 2.7 kat daha hızlı olduğu görülmüştür. Doğruluk oranlarının karşılaştırılmasında ise Nvidia Jetson TK1'in biraz daha yüksek doğruluk oranlarına sahip olduğu görülmüştür. Sonuç olarak, grafik işleme kartları kullanılarak elde edilen hız ve doğruluk değerleri, bu gömülü sistemlerin gerçek zamanlı uygulamalarda başarıyla kullanılabileceğini kanıtlanmıştır.



Şekil 7. LeNet ağı blok diyagramı (shadowthink websitesi)

Sonuçlar

Bu çalışmada, derin öğrenmenin bir çeşidi olan ve özellikle görüntü işleme uygulamalarında çok kullanılan DKSA'ların katmanları kısaca tanıtılmış ve mimarisi verilmiştir. DKSA'ları uygulamak için Caffe programı kullanılmıştır. Uygulamalar gömülü olarak Nvidia Jetson TK1/TX1 geliştirme kartları ve Nvidia GTX550/ GTX960 ekran kartlarına sahip iki bilgisayar üzerinde gerçekleştirilmiştir. Kartlar ve bilgisayar üzerinde, LeNet ağı ile el yazısı rakamları tanıma yapılmıştır. Başarım, hız ve doğruluk açısından değerlendirilmiştir. Elde edilen sonuçlar Tablo 3'de gösterilmiştir. Tablo 3'deki sonuçlar, GPU'ların CPU'lara göre doğruluk oranlarının yakın olmasına karşın daha hızlı olduğunu göstermiştir.

Kaynaklar

LeCun, Y., Boser, B., Denker, J.S., Henderson, D.R., Howard, E., Hubbard, W., and Jackel, L.D., (1989). "Backpropagation applied to handwritten zip code recognition", Neural Computation, 1, 4, 541–551.

Bengio, Y., (2009). "Learning deep architectures for AI. Foundations and Trends in Machine Learning", 1, 1–127.

Uçar, A., Demir, Y., and Güzeliş, C., (2016). "Moving Towards in Object Recognition with Deep Learning for Autonomous Driving Applications", INnovations in Intelligent SysTems and Applications (INISTA), International Symposium on IEEE.

Deng, L., and Yu, D., (2014). "Deep Learning: Methods and Applications", Foundations and Trends in Signal Processing, 7, 3–4.

Christian, S., Toshev, A., and Ethan, D., (2013). "Deep neural networks for object detection", Advances in Neural Information Processing Systems, 2553-2561.

Haykin S., (2008) "Neural Networks and Learning Machines". 3rd ed. New Jersey, USA: Prentice Hall.

http://www.nvidia.com.tr/object/jetson-tk1-embedded-dev-kit-tr.html

Bauman, B., (2015). "Jetson TK1", Department of Computer Engineering University of Heidelberg Germany.

Jia,Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., (2014). "Caffe: Convolutional Architecture for Fast Feature Embedding", arXiv preprint arXiv:1408.5093.

LeCun, Y., Cortes. C., (1998). "The MNIST database of handwritten digits".

LeCun, Y., Jackel. L.D., Bottou. L., Cortes. C., Denker. J.S., Drucker. H., Guyon. I., Muller, (1995). "Learnıng algorithms for classification: a comparison on handwritten digit recognition", website:http://yann.lecun.com/exdb/publis/pdf/lecun-95a.pdf.

https://shadowthink.com/blog/tech/2016/08/28/Caffe-MNIST-tutorial

Evaluation of deep learning at GPU boards by using Caffe

Extended abstract

Artificial Neural Networks (ANNs) are used in many applications ranging from computer vision to speech recognition. These algorithms give superior results than conventional calculation methods. However, in applications where the number of data is large, it requires a lot of computation. For this reason, the use of Graphics Processing Units (GPUs), which speed up data processing by parallel processing of data, has become mandatory. In recent years, deep learning algorithms, a kind of ANNs, have been successfully implemented in many real life applications thanks to GPUs, and demand for embedded systems containing GPUs has increased.

In this study, Deep Convolutional Neural Networks (DCNNs), which are one of the deep learning methods, were briefly introduced and the low cost Nvidia Jetson TK1/TX1 embedded systems were examined in general way. The Caffe program was used to implement the deep learning algorithms. Finally, the LeNet network was trained using MNIST data on Nvidia Jetson TK1/TX1 boards with parallel processing power. CPU and GPU modes of the boards were used for training. Moreover, two computers with i7 processor were used for computing with the Nvidia Jetson TK1/TX1 boards. The results were evaluated in terms of time and accuracy.

Caffe is an open source platform developed by the University of California, Berkeley to apply deep learning algorithms. Codes are written in C ++ with the CUDA library used for GPU computation. It also supports the CUDA library, Python/Numpy, and MATLAB.

This study used two computers with Nvidia GTX550 and GTX960 graphics cards, and Nvidia Jetson TK1 / TX1 development cards. Handwriting recognition was done as an application. The MNIST handwritten dataset and LeNet network were used.

The MNIST dataset contains 60000 training images and 10000 test images in 28x28 handwritten numbers. MNIST dataset consisting of 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 handwritten digits.

LeNet is the first network structure in which convolutional networks are successfully applied. LeNet was developed by Yann LeCun in 1990. LeNet architecture is used for other applications, but it is generally used for numbers.

In the experiments, the handwritten numbers were recognized by using the LeNet network. The performances of both two boards, and the computer were evaluated in terms of speed and accuracy. The results showed that GPUs were faster than CPUs although the accuracy rates were close to each other.

Keywords: *Deep learning, convolutional neural network, LeNet, MNIST, Caffe, Nvidia Jetson TK1/TX1*