

**İTÜ**  
**LİSANSÜSTÜ DERS KATALOG FORMU**  
**(GRADUATE COURSE CATALOGUE FORM)**

<b>Dersin Adı</b>		<b>Course Name</b>		
CUDA ile GPU'lar için Paralel Hesaplama		Parallel Computing for GPUs using CUDA		
<b>Kodu (Code)</b>	<b>Yarıyıl (Semester)</b>	<b>Kredisi (Local Credits)</b>	<b>AKTS Kredisi (ECTS Credits)</b>	<b>Ders Seviyesi (Course Level)</b>
BLG5XXE	Bahar (Spring)	3	7.5	Yuksek Lisans (M.Sc.)
<b>Lisansüstü Program (Graduate Program)</b>	Computer Engineering			
<b>Dersin Türü (Course Type)</b>	Seçmeli (Elective)	<b>Dersin Dili (Course Language)</b>	İngilizce (English)	
<b>Dersin İçeriği (Course Description)</b> <i>30-60 kelime arası</i>	GPU'ların teknolojik gelişimi. Paralel algoritma tasarımına yönelik temel kriterler. CUDA programlama modeli. GPU mikro-mimarisi. Doluluk, GPU performansı, performans analizi ve hata ayıklamama araçları. Program kontrol akışına yönelik optimizasyonlar. GPU bellek sistemi. Bellek performansını iyileştirmeye yönelik optimizasyonlar. Bütünleşik bellek ve CPU-GPU programlama. Senkronizasyon, atomik işlemler, bellek tutarlılığına yönelik teknikler. Paralel şablonlar (histogram, prefix sum, vs.). Dinamik paralellik ve çoklu-GPU programlama. CUDA kütüphaneleri (CuBlas, CuDNN, vs.). Diğer GPU programlama ortamları, OpenCL, HSA.  History of GPUs. Principles of parallel algorithm design. CUDA programming model. GPU micro-architecture. Occupancy, GPU performance and performance analysis and debugging tools. Branch divergence and control flow optimizations. GPU memory system. Memory system optimizations. Unified Memory and CPU-GPU cooperative computing. Synchronization, atomics, cache coherence, memory consistency. Parallel patterns, Histogram, Prefix-sum, etc. Dynamic parallelism, streams, multi-GPU programming. CUDA libraries, CuBlas, CuDNN, etc. Other GPU programming environments, OpenCL and HSA.			
<b>Dersin Amacı (Course Objectives)</b> <i>Maddeler halinde 2-5 adet</i>	CUDA programlama modelini ve GPU mikro-mimarisini anlamak. CUDA programlarının performans etkenlerini anlamak ve verimli çalışacak CUDA programları geliştirmeyi öğrenmek.  Understanding the CUDA programming model and GPU micro-architecture. Studying and understanding the performance of CUDA programs and learning to develop efficient parallel programs using CUDA.			
<b>Dersin Öğrenme Çıktıları (Course Learning Outcomes)</b> <i>Maddeler halinde 4-9 adet</i>	<ol style="list-style-type: none"><li>1. Paralel algoritma tasarımına yönelik temel kriterleri listeleyebilme ve kullanabilme.</li><li>2. Yaygın kullanılan GPU mimarilerini ve programlama modellerini anlatabilme.</li><li>3. GPU programlarının performansını etkileyen temel faktörleri listeleyebilme.</li><li>4. Program kontrol akışına dayalı temel optimizasyonları açıklayabilme ve uygulayabilme.</li><li>5. Bellek performansını iyileştirmeye yönelik optimizasyonları açıklayabilme ve uygulayabilme.</li><li>6. GPU programlarında veri alışverişinin doğru sağlanmasına, senkronizasyona yönelik teknikleri öğrenme, uygulama ve performansa getirdiği yükü anlatabilme.</li><li>7. Paralel şablonların CUDA ile gerçekleştirimlerini kullanabilme.</li></ol> <ol style="list-style-type: none"><li>1. Listing and applying fundamental principles of parallel algorithm design.</li><li>2. Describe common GPU architectures and programming models.</li><li>3. Explaining the factors that affect the performance of GPU programs.</li><li>4. Describing and applying the fundamental control flow optimizations to GPU programs.</li><li>5. Describing and applying the fundamental memory optimizations to GPU programs.</li><li>6. Explaining the techniques for communication and synchronization in GPU and understanding the overhead of communication on GPUs.</li><li>7. Using parallel patterns with CUDA.</li></ol>			

<b>Kaynaklar</b> (References) <u>En önemli 5 adedini belirtiniz.</u>	<ul style="list-style-type: none"> <li>David B. Kirk and Wen-mei Hwu, 2012, Programming Massively Parallel Processors: A Hands-on Approach, (2nd Edition), Morgan Kaufmann, ISBN 0-12-415992-3.</li> <li>Gram, A. Gupta, G. Karypis, and V. Kumar, Introduction to Parallel Computing, 2003, (2nd Ed.), Addison-Wesley.</li> <li>NVidia, <i>CUDA Programming Guide</i>, Available from <a href="http://www.nvidia.com/object/cuda_develop.html">http://www.nvidia.com/object/cuda_develop.html</a> for CUDA 2.0 and Windows, Linux or MAC OS.</li> </ul>		
<b>Ödevler ve Projeler</b> (Homework & Projects)	3 ödev (programlama ve performans analizi)		
<b>Laboratuar Uygulamaları</b> (Laboratory Work)	3 homeworks (programming and performance analysis study)		
<b>Bilgisayar Kullanımı</b> (Computer Use)			
<b>Diğer Uygulamalar</b> (Other Activities)			
<b>Başarı Değerlendirme Sistemi</b> (Assessment Criteria)	<b>Faaliyetler</b> (Activities)	<b>Adedi*</b> (Quantity)	<b>Değerlendirmedeki Katkısı, %</b> (Effects on Grading, %)
	Yıl İçi Sınavları (Midterm Exams)		
	Kısa Sınavlar (Quizzes)		
	Ödevler (Homework)	3	%60
	Projeler (Projects)		
	Dönem Ödevi/Projesi (Term Paper/Project)		
	Laboratuar Uygulaması (Laboratory Work)		
	Diğer Uygulamalar (Other Activities)		
	Final Sınavı (Final Exam)	1	%40

\*Yukarıda Belirtilen Sayılar Minimum Olup Yerine Getirilmesi Zorunludur.

## DERS PLANI

Hafta	Konular	Dersin Çıktıları
1	GPU'ların teknolojik gelişimi ve GPU'ların genel amaçlı programların hızlandırılmaları için kullanılmaları.	1
2	Paralel algoritma tasarımına yönelik temel kriterler.	1,2,3
3	CUDA programlama modeli ve basit bir CUDA programının geliştirilmesi.	1,2,3
4	GPU mikro-mimarisine bakış.	1,2,3
5	GPU'larda doluluk, CUDA programlarının performansı ve performans analizi için araçlar.	1,2,3
6	Dallanma ve program akışına yönelik optimizasyonlar.	3,4
7	GPU bellek sistemi ve bellek performansını iyileştirmeye yönelik optimizasyonlar.	3,5
8	Bütünleşik bellek, CPU-GPU programlama.	1,2,3,5
9	Senkronizasyon, önbellekte veri bütünlüğü, atomik işlemler, bellek tutarlılığı.	2,6
10	Paralel şablonlar (Histogram, Prefix-sum, Scan, Reduction).	3,7
11	Dinamik paralellik, çoklu-GPU programlama.	1,2
12	CUDA kütüphaneleri, cuBLAS, cuDNN.	1,2,3,4,5,6,7
13	GPU'larda çizge hesaplama ve data analitik algoritmaları	1,2,3,4,5,6,7
14	Diğer genel amaçlı GPU hesaplama ortamları, OpenCL, HSA.	1,2,3,4,5,6,7

## COURSE PLAN

Weeks	Topics	Course Outcomes
1	History of GPUs, introduction to General Purpose GPU computing.	1
2	Principles of parallel algorithm design.	1,2,3
3	CUDA programming model and writing a simple GPU program.	1,2,3
4	GPU micro-architecture.	1,2,3
5	Occupancy, Performance of CUDA programs, performance tools.	1,2,3
6	Branch divergence and control flow optimizations.	3,4
7	GPU memory system and optimizing memory.	3,5
8	Unified Memory and CPU-GPU cooperative computing.	1,2,3,5
9	Cache coherence, synchronization, atomics, and memory consistency on GPUs.	2,6
10	Parallel patterns, Histograms, Prefix sum, Scan, Reduction.	3,7
11	Dynamic parallelism, streams, multi-GPU programming.	1,2
12	CUDA libraries, cuBLAS, cuDNN, etc.	1,2,3,4,5,6,7
13	Graph computing and data analytics with GPUs.	1,2,3,4,5,6,7
14	Other GPGPU computing environments, OpenCL, HSA.	1,2,3,4,5,6,7

**NOT-1: Ders planı, sadece hafta bazında işlenen ders konularını içermeli, ara ve kısa sınavlar ders planlarına yazılmamalıdır.**

### Dersin Bilgisayar Mühendisliği Doktora Programıyla İlişkisi

Programın mezuna kazandıracığı bilgi, beceri ve yetkinlikler (programa ait çıktılar)	Katkı Seviyesi		
	1	2	3

i.	Lisans düzeyi yeterliliklerine dayalı olarak, ilgili program alanında bilgilerini uzmanlık düzeyinde geliştirebilme ve derinleştirebilme (yeterli bilgi birikimi) (bilgi).		X	
ii.	Alanının ilişkili olduğu disiplinler arası etkileşimi kavrayabilme (bilgi).			
iii.	Alanında edindiği uzmanlık düzeyindeki kuramsal ve uygulamalı bilgileri kullanabilme (beceri).			X
iv.	Alanında edindiği bilgileri farklı disiplin alanlarından gelen bilgilerle bütünleştirerek yorumlayabilme ve yeni bilgiler oluşturabilme (beceri).			
v.	Alanını ile ilgili karşılaşılan sorunları araştırma yöntemlerini kullanarak çözümlenebilir (beceri).			X
vi.	Alanını ile ilgili uzmanlık gerektiren bir çalışmayı bağımsız olarak yürütebilme (Bağımsız Çalışabilme ve Sorumluluk Alabilme Yetkinliği).		X	
vii.	Alanı ile ilgili uygulamalarda karşılaşılan ve öngörülemeyen karmaşık sorunların çözümü için yeni stratejik yaklaşımlar geliştirebilme ve sorumluluk alarak çözüm üretebilme (Bağımsız Çalışabilme ve Sorumluluk Alabilme Yetkinliği).	X		
viii.	Alanı ile ilgili sorunların çözümlenmesini gerektiren ortamlarda liderlik yapabilme (Bağımsız Çalışabilme ve Sorumluluk Alabilme Yetkinliği).			
ix.	Alanında edindiği uzmanlık düzeyindeki bilgi ve becerileri eleştirel bir yaklaşımla değerlendirebilme ve öğrenmesini yönlendirebilme (Öğrenme Yetkinliği).		X	
x.	Alanındaki güncel gelişmeleri ve kendi çalışmalarını, nicel ve nitel veriler ile destekleyerek, alanındaki ve alan dışındaki gruplara, yazılı, sözlü ve görsel olarak sistemli biçimde aktarabilme (İletişim ve Sosyal Yetkinlik).		X	
xi.	Sosyal ilişkileri ve bu ilişkileri yönlendiren normları eleştirel bir bakış açısı ile inceleyebilme, geliştirebilme ve gerektiğinde değiştirmek üzere harekete geçebilme (İletişim ve Sosyal Yetkinlik).			
xii.	Alanının gerektirdiği düzeyde bilgisayar yazılımı ile birlikte bilişim ve iletişim teknolojilerini ileri düzeyde kullanabilme (İletişim ve Sosyal Yetkinlik).			
xiii.	Alanı ile ilgili verilerin toplanması, yorumlanması, uygulanması ve duyurulması aşamalarında toplumsal, bilimsel, kültürel ve etik değerleri gözetenek denetleyebilme ve bu değerleri öğretebilme (Alana Özgü Yetkinlik).			
xiv.	Alanı ile ilgili konularda strateji, politika ve uygulama planları geliştirebilme ve elde edilen sonuçları, kalite süreçleri çerçevesinde değerlendirebilme (Alana Özgü Yetkinlik).			
xv.	Alanında özümstedikleri bilgiyi, problem çözme ve/veya uygulama becerilerini, disiplinlerarası çalışmalarda kullanabilme (Alana Özgü Yetkinlik).			
xvi.	Tezli programlarda, kendi çalışmalarını, alanındaki uluslararası platformlarda, yazılı, sözlü ve/veya görsel olarak aktarabilme (Alana özgü yetkinlik).			

1: Az, 2. Kısmi, 3. Tam

## Relationship between the Course and Computer Engineering PhD. Program

X	Program Outcomes	Level of Contribution		
		1	2	3
i.	Developing and intensifying knowledge in the related program's area, based upon the competency in the undergraduate level (sufficient knowledge) (knowledge).		X	
ii.	Grasping the inter-disciplinary interaction related to one's area (knowledge).			
iii.	The ability to use the expert-level theoretical and practical knowledge acquired in the area(skill).			X
iv.	Interpreting and forming new types of knowledge by combining the knowledge from the area and the knowledge from various other disciplines (skill).			
v.	Solving the problems faced in the area by making use of the research methods (skill).			X
vi.	The ability to carry out a specialistic study related to one's area independently. (Competence to work independently and take responsibility).		X	
vii.	Developing new strategic approaches to solve the unforeseen and complex problems arising in the practical processes of one's area and coming up with solutions while taking responsibility (Competence to work independently and take responsibility).	X		
viii.	Fulfilling the leader role in the environments where solutions are sought for the problems related to the area (Competence to work independently and take responsibility)			
ix.	Assessing the specialistic knowledge and skill gained through the study with a critical view and directing one's own learning process (Learning Competence).		X	
x.	Systematically transferring the current developments in the area and one's own work to other groups in and out of the area; in written, oral and visual forms (Communication and Social Competency).		X	
xi.	Ability to see and develop social relationships and the norms directing these relationships with a critical look and the ability to take action to change these when necessary. (Communication and Social Competency).			
xii.	Using the computer software together with the information and communication technologies efficiently and according to the needs of the area (Communication and Social Competency).			
xiii.	Paying regard to social, scientific, cultural and ethical values during the collecting, interpreting, practicing and announcing processes of the area related data and the ability to teach these values to others (Area Specific Competency).			
xiv.	Developing strategy, policy and application plans concerning the subjects related to the area and the ability to evaluate the end results of these plans within the frame of quality processes (Area Specific Competency).			
xv.	Using the knowledge and the skills for problem solving and/or application (which are processed within the area) in inter-disciplinary studies (Area Specific Competency).			
xvi.	In the programs with thesis, the ability to present one's own work within the international environments orally, visually and in written forms (Area Specific Competency).			

1: Little, 2. Partial, 3. Full

**NOT-2: Ders ile ilgisi olmayan çıktıların boş bırakılması gerekmektedir.**

<u>Düzenleyen (Prepared by)</u> Ayşe Yılmaz	<u>Tarih (Date)</u> 31/10/2018	<u>İmza (Signature)</u>
--	-----------------------------------	-------------------------