

İNTELLEKTUAL HESABLAMA PARADİQMASI VƏ MÜASİR KOMPÜTER ARXİTEKTURALARININ TƏKAMÜLÜ

NAĞİYEVA ZÜMRÜD SAHİBQULU QIZI

Azərbaycan Dövlət Neft və Sənaye Univesitetinin nəzdində
Sənaye və Texnologiya kolleci, müəllim

Xülasə. Müasir informasiya texnologiyalarının sürətli inkişafı və süni intellekt sistemlərinin geniş tətbiqi kompüter arxitekturalarının yeni tələblərə uyğun şəkildə yenidən formalaşdırılmasını zəruri etmişdir. Ənənəvi hesablama sistemləri böyük həcmli verilənlərin emalı, maşın təlimi və dərin öyrənmə alqoritmlərinin icrası zamanı müəyyən məhdudiyyətlərlə üzləşdiyindən, intellektual hesablama paradigması yeni arxitektura yanaşmaların meydana çıxmasına səbəb olmuşdur. Məqalədə intellektual hesablama paradigmasının mahiyyəti, onun əsas xüsusiyyətləri və müasir kompüter arxitekturalarının inkişafına təsiri araşdırılmışdır. Həmçinin CPU, GPU, TPU və NPU kimi ixtisaslaşmış hesablama qurğularının xüsusiyyətləri təhlil edilmiş, heterogen hesablama sistemlərinin və süni intellekt yönümlü arxitekturaların üstünlükləri göstərilmişdir. Tədqiqat nəticəsində müəyyən edilmişdir ki, süni intellekt tətbiqlərinin artan hesablama tələbləri kompüter arxitekturalarında paralellik, enerji səmərəliliyi və yüksək məhsuldarlıq prinsiplərinin daha geniş tətbiqinə şərait yaradır. Məqalədə həmçinin gələcəyin hesablama sistemlərinin inkişaf perspektivləri və intellektual texnologiyaların bu sahədə yaratdığı yeni imkanlar qiymətləndirilmişdir.

Açar sözlər: intellektual hesablama, süni intellekt, kompüter arxitekturası, GPU, TPU, NPU, heterogen hesablama sistemləri, paralel emal, yüksək məhsuldarlıq.

THE INTELLECTUAL COMPUTING PARADIGM AND THE EVOLUTION OF MODERN COMPUTER ARCHITECTURES

Summary. The rapid development of modern information technologies and the widespread use of artificial intelligence systems have necessitated the reformulation of computer architectures in accordance with new requirements. Since traditional computing systems face certain limitations when processing large amounts of data, executing machine learning and deep learning algorithms, the intelligent computing paradigm has led to the emergence of new architectural approaches. The article examines the essence of the intelligent computing paradigm, its main features and its impact on the development of modern computer architectures. The characteristics of specialized computing devices such as CPU, GPU, TPU and NPU are also analyzed, and the advantages of heterogeneous computing systems and artificial intelligence-oriented architectures are shown. As a result of the study, it was determined that the increasing computational requirements of artificial intelligence applications create conditions for the wider application of the principles of parallelism, energy efficiency and high performance in computer architectures. The article also evaluates the development prospects of future computing systems and the new opportunities created by intelligent technologies in this area.

Keywords: intelligent computing, artificial intelligence, computer architecture, GPU, TPU, NPU, heterogeneous computing systems, parallel processing, high performance.

Резюме. Быстрое развитие современных информационных технологий и широкое использование систем искусственного интеллекта обусловили необходимость переосмысления компьютерных архитектур в соответствии с новыми требованиями. Поскольку традиционные вычислительные системы сталкиваются с определенными ограничениями при обработке больших объемов данных, выполнении алгоритмов машинного обучения и глубокого обучения, парадигма интеллектуальных вычислений привела к появлению новых архитектурных подходов. В статье рассматривается сущность

ОФ "Международный научно-исследовательский центр "Endless Light in Science"

парадигмы интеллектуальных вычислений, ее основные особенности и влияние на развитие современных компьютерных архитектур. Также анализируются характеристики специализированных вычислительных устройств, таких как ЦП, ГП, ТПУ и НПУ, и показаны преимущества гетерогенных вычислительных систем и архитектур, ориентированных на искусственный интеллект. В результате исследования установлено, что растущие вычислительные требования приложений искусственного интеллекта создают условия для более широкого применения принципов параллелизма, энергоэффективности и высокой производительности в компьютерных архитектурах. В статье также оцениваются перспективы развития будущих вычислительных систем и новые возможности, создаваемые интеллектуальными технологиями в этой области.

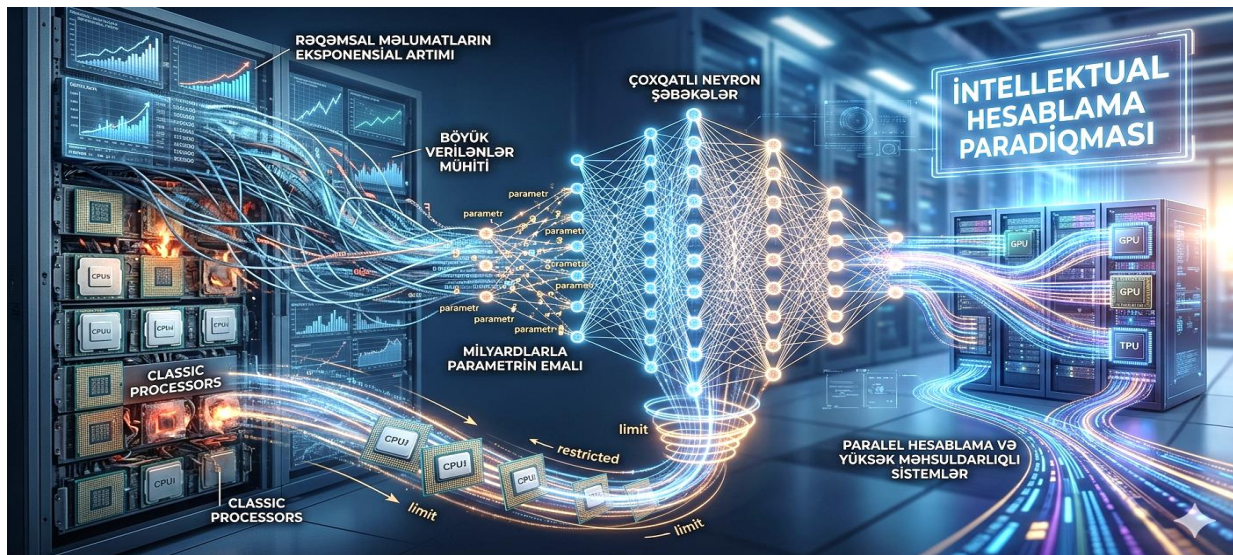
Ключевые слова: интеллектуальные вычисления, искусственный интеллект, компьютерная архитектура, GPU, TPU, NPU, гетерогенные вычислительные системы, параллельная обработка, высокая производительность.

XXI əsrin ikinci onilliyindən etibarən süni intellekt texnologiyalarının sürətli inkişafı hesablama sistemlərinə olan tələbləri köklü şəkildə dəyişdirmişdir. Böyük həcmli verilənlərin emalı, maşın öyrənməsi modellərinin təlimi, dərin neyron şəbəkələrinin qurulması və generativ süni intellekt tətbiqlərinin geniş yayılması ənənəvi hesablama sistemlərinin imkanlarını məhdudlaşdıran amillər kimi çıxış etmişdir. Xüsusilə böyük dil modellərinin, kompüter görməsi sistemlərinin və avtonom qərarvermə mexanizmlərinin inkişafı yüksək məhsuldarlığa malik hesablama platformalarına olan ehtiyacı artırmışdır. Bu şəraitdə kompüter arxitekturası sahəsində yeni yanaşmaların formalaşdırılması və mövcud arxitekturaların modernləşdirilməsi zərurəti meydana çıxmışdır. Ənənəvi J.F. Neumann arxitekturası uzun müddət kompüter sistemlərinin əsasını təşkil etsə də, verilənlərin emalı və yaddaş arasında məlumat mübadiləsinin yaratdığı gecikmələr, enerji sərfiyyatının artması və paralel hesablama imkanlarının məhdudluğu müasir süni intellekt tətbiqlərinin tələblərini tam şəkildə qarşılamağa imkan vermir. Nəticədə heterogen hesablama sistemləri, qrafik prosessorlar, tensor emal blokları və neyron emal qurğuları kimi yeni arxitektur həllər meydana çıxmışdır. Müasir dövrdə intellektual hesablama paradiqması yalnız hesablama məhsuldarlığının artırılmasını deyil, həm də sistemlərin adaptivliyini, enerji səmərəliliyini və böyük verilənlərlə işləmə qabiliyyətini təmin edən əsas istiqamətlərdən biri hesab olunur. Bu baxımdan kompüter arxitekturalarının təkamülünün süni intellekt texnologiyalarının inkişafı ilə qarşılıqlı əlaqədə öyrənilməsi mühüm elmi və praktiki əhəmiyyət kəsb edir.

İntellektual hesablama paradiqması kompüter sistemlərinin yalnız verilmiş əmrləri yerinə yetirən klassik hesablama mexanizmləri kimi deyil, həm də verilənləri təhlil edən, öyrənən, qərar qəbul edən və öz fəaliyyətini optimallaşdıran sistemlər kimi fəaliyyət göstərməsini nəzərdə tutur. Bu paradiqmanın əsas məqsədi insan intellektinə xas olan bəzi xüsusiyyətlərin kompüter sistemlərində modelləşdirilməsi və tətbiq edilməsidir. Müasir elmi ədəbiyyatda intellektual hesablama anlayışı süni intellekt, maşın öyrənməsi, dərin öyrənmə, qeyri-səlis məntiq, ekspert sistemləri və təkamül alqoritmləri kimi texnologiyaların integrasiyası əsasında formalaşan kompleks hesablama yanaşması kimi xarakterizə edilir. İntellektual hesablama paradiqmasının meydana gəlməsi bir sıra texnoloji və elmi amillərlə bağlıdır. İlk növbədə, rəqəmsal məlumatların eksponensial artımı ənənəvi hesablama üsullarının effektivliyini azaltmışdır. Böyük verilənlər mühitində informasiya axınlarının operativ şəkildə emalı üçün paralel hesablama və yüksək məhsuldarlıqlı sistemlərə ehtiyac yaranmışdır. Digər tərəfdən, süni intellekt alqoritmlərinin mürəkkəbləşməsi daha böyük hesablama resurslarının tələb olunmasına səbəb olmuşdur. Xüsusilə çoxqatlı neyron şəbəkələrinin təlim prosesində milyardlarla parametrlərin emal edilməsi klassik prosessorların imkanlarını məhdudlaşdırmış və yeni arxitektur həllərin meydana çıxmasını sürətləndirmişdir.

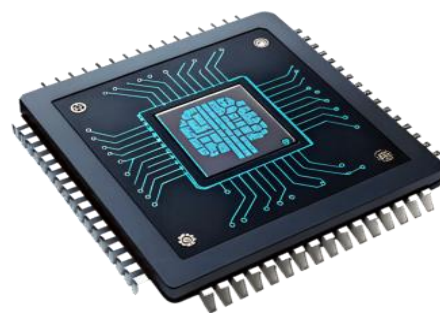
İntellektual hesablama paradiqmasının əsas xüsusiyyətləri adaptivlik, paralellik, özünüöyrənmə qabiliyyəti, yüksək məhsuldarlıq və enerji səmərəliliyidir. Bu xüsusiyyətlər ənənəvi alqoritmik yanaşmalardan fərqli olaraq sistemlərin dəyişən mühit şəraitinə uyğunlaşmasına və

qərarvermə prosesinin optimallaşdırılmasına imkan yaradır. Müasir süni intellekt sistemlərində istifadə olunan dərin neyron şəbəkələri, transformator modelləri və generativ arxitekturalar məhz bu paradigmanın praktik tətbiqləri hesab olunur. Bununla yanaşı, intellektual hesablama yanaşması kompüter arxitekturasının inkişafına birbaşa təsir göstərərək CPU mərkəzli sistemlərdən heterogen və ixtisaslaşdırılmış hesablama platformalarına keçidi sürətləndirmişdir.



Kompüter texnologiyalarının inkişaf tarixi hesablama sistemlərinin arxitekturasında baş verən təkamül prosesləri ilə sıx bağlıdır. İlk elektron hesablama maşınlarının yaradılmasından başlayaraq müasir yüksək məhsuldarlıqlı sistemlərə qədər kompüter arxitekturaları daim təkmilləşdirilmiş və yeni texnoloji tələblərə uyğunlaşdırılmışdır. Bu inkişaf prosesinin əsasını XX əsrin ortalarında irəli sürülmüş Neumann arxitekturası təşkil etmişdir. Program və verilənlərin eyni yaddaş sahəsində saxlanması prinsipinə əsaslanan bu model uzun illər ərzində kompüter sistemlərinin layihələndirilməsində fundamental konsepsiya kimi istifadə olunmuşdur. Neumann tərəfindən təklif edilən mərkəzi prosessor, operativ yaddaş, giriş-çıkış qurğuları və məlumat ötürmə magistrallarından ibarət struktur kompüter texnologiyalarının geniş yayılmasına və standartlaşmasına mühüm töhfə vermişdir.

Ənənəvi arxitekturaların inkişaf mərhələləri bir neçə əsas dövrə bölünə bilər. Birinci mərhələ vakuum lampaları əsasında qurulan ilkin elektron hesablama sistemlərini əhatə edir. Sonrakı mərhələdə tranzistorların və integral sxemlərin tətbiqi kompüterlərin ölçülərinin kiçilməsinə, etibarlılığının və hesablama sürətinin artmasına səbəb olmuşdur. Mikroprosessor texnologiyalarının yaranması isə fərdi kompüterlərin inkişafına təkan vermişdir. Daha sonra çoxnüvəli prosessorların meydana çıxması paralel emal imkanlarını genişləndirmiş və yüksək məhsuldarlığa malik hesablama sistemlərinin formalaşmasına şərait yaratmışdır. Bununla belə, hesablama gücünün artmasına baxmayaraq, ənənəvi arxitekturalar süni intellekt və böyük verilənlər texnologiyalarının tələbləri qarşısında müəyyən məhdudiyyətlərlə üzləşmişdir. Ənənəvi kompüter arxitekturalarının digər məhdudiyyətlərindən biri onların süni intellekt tətbiqlərində tələb olunan yüksək səviyyəli paralelliyi effektiv şəkildə dəstəkləməməsidir. Klassik CPU-lar ardıcıl əməllərin icrası üçün optimallaşdırılmışdır və mürəkkəb riyazi əməliyyatların kütləvi şəkildə paralel emalı zamanı səmərəliliyi azalır. Dərin öyrənmə modellərinin təlim prosesində isə milyonlarla matris əməliyyatlarının eyni vaxtda icra olunması tələb olunur. Bu səbəbdən ənənəvi prosessorlar süni intellekt yüklərinin emalında kifayət qədər yüksək performans göstərə bilmir. Paralel emalın məhdudluğu, enerji sərfiyyatının yüksək olması



və yaddaş-prosessor qarşılıqlı əlaqəsində yaranan gecikmələr yeni arxitektura yanaşmalarının meydana çıxmasını zəruri etmişdir.

Son onilliklər ərzində süni intellekt texnologiyalarının sürətli inkişafı kompüter sistemlərinə olan tələbləri əsaslı şəkildə dəyişdirmiş və kompüter arxitekturalarının yeni istiqamətlər üzrə inkişafını stimullaşdırmışdır. Ənənəvi proqram təminatları əsasən ardıcıl hesablama əməliyyatlarına əsaslandığı halda, müasir süni intellekt tətbiqləri böyük həcmli verilənlərin emalı, çoxsaylı matris hesablamaları və paralel işləmə mexanizmlərini tələb edir. Xüsusilə maşın öyrənməsi və dərin öyrənmə alqoritmlərinin geniş yayılması nəticəsində hesablama sistemlərinin məhsuldarlığı, yaddaş tutumu və enerji səmərəliliyi əsas prioritetlərə çevrilmişdir. Süni intellekt tətbiqlərinin hesablama mürəkkəbliyi artdıqca klassik CPU əsaslı sistemlərin imkanları kifayət etməmiş və yeni nəsil arxitektura həllərin hazırlanması zərurəti yaranmışdır.

Süni intellektin kompüter arxitekturasına təsiri ilk növbədə hesablama proseslərinin təşkilində özünü göstərir. Dərin neyron şəbəkələrinin təlimi və inferensiya mərhələləri milyonlarla, bəzən isə milyardlarla parametrlərin emal edilməsini tələb edir. Bu proseslər zamanı matris-vurulma əməliyyatları, vektor hesablamaları və paralel verilən emalı xüsusi əhəmiyyət kəsb edir. Ənənəvi prosessorlar mürəkkəb məntiqi əməliyyatların yerinə yetirilməsi üçün optimallaşdırıldığı halda, süni intellekt modellərinin tələblərinə uyğun olaraq minlərlə əməliyyatın eyni vaxtda icrası daha vacib hesab olunur. Bu səbəbdən paralel hesablama prinsiplərini dəstəkləyən yeni prosessor arxitekturaları meydana çıxmışdır.

Süni intellekt texnologiyalarının inkişafı yaddaş sistemlərinin arxitekturasına da ciddi təsir göstərmişdir. Müasir neyron şəbəkələri böyük həcmli verilənlərin operativ şəkildə saxlanılmasını və emal olunmasını tələb etdiyindən yüksək ötürmə qabiliyyətinə malik yaddaş texnologiyalarına ehtiyac yaranmışdır. HBM (High Bandwidth Memory) kimi yeni nəsil yaddaş həlləri məhz bu tələblərin ödənilməsi məqsədilə hazırlanmışdır. Bununla yanaşı, hesablama qurğularının yaddaşa daha yaxın yerləşdirilməsi prinsipinə əsaslanan “Processing-in-Memory” və “Near-Memory Computing” yanaşmaları Von Neumann darboğazının azaldılmasına yönəlmiş mühüm texnoloji istiqamətlərdən hesab edilir.

Cədvəl 1. Süni intellektin təsiri nəticəsində kompüter arxitekturasında baş verən əsas dəyişikliklər

Arxitektura komponent	Ənənəvi yanaşma	Müasir AI-yönümlü yanaşma
Prosesor	CPU əsaslı hesablama	CPU + GPU + TPU + NPU
Hesablama modeli	Ardıcıl emal	Kütləvi paralel emal
Yaddaş sistemi	DRAM əsaslı yaddaş	HBM, PIM, Near-Memory Computing
Performans meyarı	Saat tezliyi	Paralellik və məhsuldarlıq
Enerji səmərəliliyi	İkinci dərəcəli amil	Əsas optimallaşdırma meyarı
Arxitektura növü	Homogen sistemlər	Heterogen sistemlər
Tətbiq sahəsi	Ümumi təyinatlı proqramlar	Süni intellekt və böyük verilənlər
Məlumat emalı	Prosesor mərkəzli	Verilən-mərkəzli hesablama

Beləliklə, süni intellekt kompüter arxitekturalarının inkişafında əsas hərəkətverici qüvvələrdən birinə çevrilmişdir. Müasir hesablama sistemlərinin layihələndirilməsi artıq yalnız ümumi təyinatlı hesablamalara deyil, həm də süni intellekt yüklərinin səmərəli emalına yönəlmişdir. Bu tendensiya gələcək illərdə daha da güclənəcək və kompüter arxitekturalarının inkişafında ixtisaslaşmış prosessorlar, heterogen sistemlər və intellektual sürətləndiricilər əsas istiqamətlərdən biri olacaqdır.

Süni intellekt texnologiyalarının sürətli inkişafı hesablama sistemlərinin arxitekturasında fundamental dəyişikliklərə səbəb olmuşdur. Xüsusilə maşın öyrənməsi, dərin öyrənmə və generativ süni intellekt modellərinin tətbiqi ənənəvi prosessorların imkanlarını aşan hesablama tələbləri formalaşdırmışdır. Bu səbəbdən müasir kompüter sistemlərində yalnız mərkəzi prosessorlardan

(CPU) istifadə edilməsi kifayət etmir. Süni intellekt yüklərinin səmərəli şəkildə emal olunması üçün qrafik prosessorlar (GPU), tensor emal qurğuları (TPU) və neyron emal blokları (NPU) kimi ixtisaslaşdırılmış prosessor arxitekturaları hazırlanmışdır. Bu qurğuların əsas məqsədi yüksək paralellik, aşağı gecikmə müddəti və enerji səmərəliliyi təmin etməklə süni intellekt alqoritmlərinin məhsuldarlığını artırmaqdır.

CPU (Central Processing Unit) uzun illər ərzində kompüter sistemlərinin əsas hesablama elementi olmuşdur. CPU-lar mürəkkəb məntiqi əməliyyatların, idarəetmə funksiyalarının və ardıcıl hesablama proseslərinin icrası üçün optimallaşdırılmışdır. Müasir çoxnüvəli CPU-lar müəyyən səviyyədə paralel emal imkanlarına malik olsa da, onların arxitekturası əsasən ümumi təyinatlı hesablama tapşırıqlarına yönəlmişdir. Süni intellekt tətbiqlərində istifadə olunan çoxsaylı matris əməliyyatları və vektor hesablamaları zamanı CPU-ların məhsuldarlığı məhdudlaşır. Buna baxmayaraq, CPU-lar süni intellekt sistemlərində idarəedici komponent kimi mühüm rol oynayır və digər sürətləndiricilərin fəaliyyətini koordinasiya edir. GPU (Graphics Processing Unit) əvvəlcə kompüter qrafikasının emalı üçün nəzərdə tutulsa da, sonradan paralel hesablama imkanları səbəbindən süni intellekt sahəsində geniş tətbiq olunmağa başlamışdır. GPU minlərlə hesablama nüvəsinə malik olduğundan böyük həcmli verilənlərin paralel şəkildə emalını təmin edir. Dərin neyron şəbəkələrinin təlimi zamanı milyonlarla riyazi əməliyyatın eyni vaxtda icra olunması tələb olduğundan GPU-lar bu sahədə yüksək səmərəlilik nümayiş etdirir. Müasir generativ süni intellekt modellərinin, o cümlədən böyük dil modellərinin hazırlanması və öyrədilməsi əsasən GPU əsaslı hesablama klasterləri vasitəsilə həyata keçirilir. GPU-ların yüksək məhsuldarlığı onların süni intellekt tədqiqatlarında əsas hesablama platformasına çevrilməsinə səbəb olmuşdur.

Son illərdə inkişaf etdirilən Neural Processing Unit (NPU) arxitekturaları isə süni intellekt funksiyalarının birbaşa son istifadəçi qurğularında həyata keçirilməsinə imkan yaradır. NPU-lar mobil telefonlar, planşetlər, ağıllı kameralar və digər kənar hesablama (edge computing) cihazlarında integrasiya olunaraq neyron şəbəkələrinin lokal şəkildə işlədilməsini təmin edir. Bu yanaşma məlumatların uzaq serverlərə göndərilməsi ehtiyacını azaldır, gecikməni minimuma endirir və məxfilik səviyyəsini artırır. Eyni zamanda NPU-lar enerji baxımından olduqca səmərəli hesab olunur və məhdud enerji resurslarına malik cihazlarda süni intellekt tətbiqlərinin reallaşdırılmasına imkan verir.

Müasir hesablama sistemlərində CPU, GPU, TPU və NPU arxitekturaları bir-birini tamamlayan komponentlər kimi fəaliyyət göstərir. CPU ümumi sistem idarəetməsini həyata keçirir, GPU böyük həcmli paralel hesablamaları yerinə yetirir, TPU dərin öyrənmə əməliyyatlarını optimallaşdırır, NPU isə süni intellekt funksiyalarını enerji baxımından səmərəli şəkildə son istifadəçi qurğularında icra edir. Bu səbəbdən müasir kompüter arxitekturalarında heterogen yanaşma üstünlük təşkil edir və müxtəlif növ prosessorların integrasiyası vasitəsilə maksimum məhsuldarlıq əldə edilir.

Cədvəl 2. AI-yönümlü prosessor arxitekturalarının müqayisəli təhlili

Xüsusiyyət	CPU	GPU	TPU	NPU
Əsas təyinat	Ümumi hesablama	Paralel hesablama	Tensor əməliyyatları	Neyron şəbəkələrinin icrası
Paralellik səviyyəsi	Aşağı-Orta	Çox yüksək	Çox yüksək	Yüksək
Enerji səmərəliliyi	Orta	Orta	Yüksək	Çox yüksək
AI təlimi üçün uyğunluq	Məhdud	Yüksək	Çox yüksək	Aşağı
AI inferensiyası üçün uyğunluq	Orta	Yüksək	Yüksək	Çox yüksək
Tətbiq sahəsi	Serverlər, fərdi kompüterlər	AI mərkəzləri, superkompüterlər	Bulud platformaları	Mobil və edge cihazlar

Hesablama məhsuldarlığı	Orta	Yüksək	Çox yüksək	Yüksək
Əsas üstünlük	<i>Universal istifadə</i>	<i>Kütləvi paralel emal</i>	<i>AI üçün optimallaşdırma</i>	<i>Aşağı enerji sərfiyyatı</i>

Nəticə olaraq:

Aparılmış tədqiqat göstərir ki, süni intellekt texnologiyalarının sürətli inkişafı kompüter arxitekturalarının təkamülündə əsas hərəkətverici amillərdən birinə çevrilmişdir. Ənənəvi Von Neumann arxitekturasının mövcud məhdudiyyətləri böyük verilənlərin emalı və süni intellekt tətbiqlərinin artan hesablama tələblərini tam şəkildə qarşılamağa imkan vermədiyindən, heterogen hesablama sistemləri və AI-yönlü prosessor arxitekturaları geniş tətbiq olunmağa başlamışdır. Tədqiqat nəticəsində müəyyən edilmişdir ki, CPU, GPU, TPU və NPU kimi ixtisaslaşdırılmış prosessorlar yüksək məhsuldarlıq, paralel emal imkanları və enerji səmərəliliyi baxımından müasir intellektual hesablama sistemlərinin əsas komponentləri hesab olunur. Həmçinin neuromorfik və kvant hesablama texnologiyaları gələcəyin kompüter arxitekturalarının formalaşmasında perspektivli istiqamətlər kimi çıxış edir. Ümumilikdə, intellektual hesablama paradigması kompüter sistemlərinin inkişafında yeni mərhələnin əsasını təşkil edir və gələcəkdə daha çevik, məhsuldar və enerji baxımından səmərəli hesablama platformalarının yaradılmasına mühüm töhfə verəcəkdir.

ƏDƏBİYYAT SİYAHISI

1. Hennessy, J. L., & Patterson, D. A. (2022). Computer architecture: A quantitative approach (7th ed.). Morgan Kaufmann.
2. Patterson, D. A., & Hennessy, J. L. (2021). Computer organization and design RISC-V edition: The hardware/software interface (2nd ed.). Morgan Kaufmann.
3. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
4. Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.
5. Jouppi, N. P., Young, C., Patil, N., Patterson, D., Agrawal, G., Bajwa, R., ... Yoon, D. H. (2017). In-datacenter performance analysis of a tensor processing unit. Proceedings of the 44th Annual International Symposium on Computer Architecture, 1–12.
6. Shalf, J. (2020). The future of computing beyond Moore's law. Philosophical Transactions of the Royal Society A, 378(2166), 20190061.
7. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.
8. Sze, V., Chen, Y. H., Yang, T. J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. Proceedings of the IEEE, 105(12), 2295–2329.
9. Mittal, S. (2019). A survey of FPGA-based accelerators for convolutional neural networks. Neural Computing and Applications, 32(4), 1109–1139.
10. Furber, S. (2016). Large-scale neuromorphic computing systems. Journal of Neural Engineering, 13(5), 051001.
11. Indiveri, G., & Liu, S. C. (2015). Memory and information processing in neuromorphic systems. Proceedings of the IEEE, 103(8), 1379–1397.
12. Preskill, J. (2018). Quantum computing in the NISQ era and beyond. Quantum, 2, 79.
13. Nielsen, M. A., & Chuang, I. L. (2010). Quantum computation and quantum information (10th anniversary ed.). Cambridge University Press.
14. Dean, J. (2021). The deep learning revolution and its implications for computer architecture and chip design. Communications of the ACM, 64(7), 58–65.
15. Pedram, M. (2020). Energy-efficient computing for machine learning and artificial intelligence. IEEE Design & Test, 37(6), 8–17.
16. Esmailzadeh, H., Blem, E., St. Amant, R., Sankaralingam, K., & Burger, D. (2012). Dark silicon and the end of multicore scaling. IEEE Micro, 32(3), 122–134.