

1n년차 게임 프로그래머가 딥러닝 칩 회사에 가서 하는 일

하재승

딥러닝?

AI가 인류를 지배할 것 이
라고 주장하는 사람들

내가 만든 신경망 모델





About

Research

Impact

Blog

Safety & Ethics

Careers



DeepMind

>

Blog

>

AlphaStar: Grandmaster level in StarCraft II using multi-agent reinforcement...



BLOG POST
RESEARCH

30 OCT 2019

AlphaStar: Grandmaster level in StarCraft II using multi-agent reinforcement learning

SHARE



AUTHORS

TL;DR: AlphaStar is the first AI to reach the top league of a widely popular esports without

딥러닝 가속 칩?

CPU → GPU → NPU/TPU

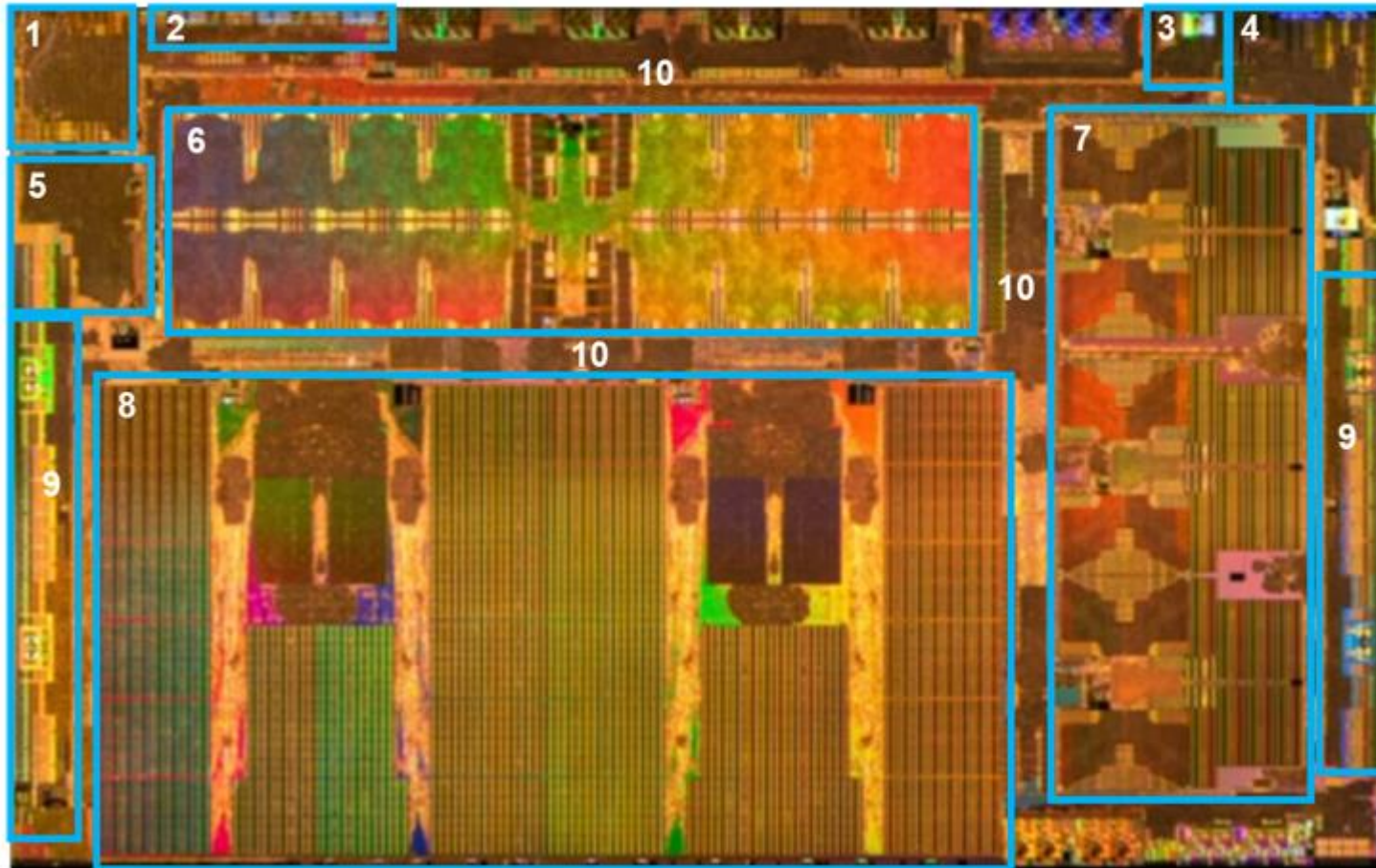
NVIDIA T4 = 200만원대

비슷한 성능의 NPU = 20~40 만원대

자율주행 (Tesla 칩)

Tesla Full Self-Driving (FSD) Chip

14nm FinFET CMOS (Samsung) – 6 Billion Transistors – 260 mm²



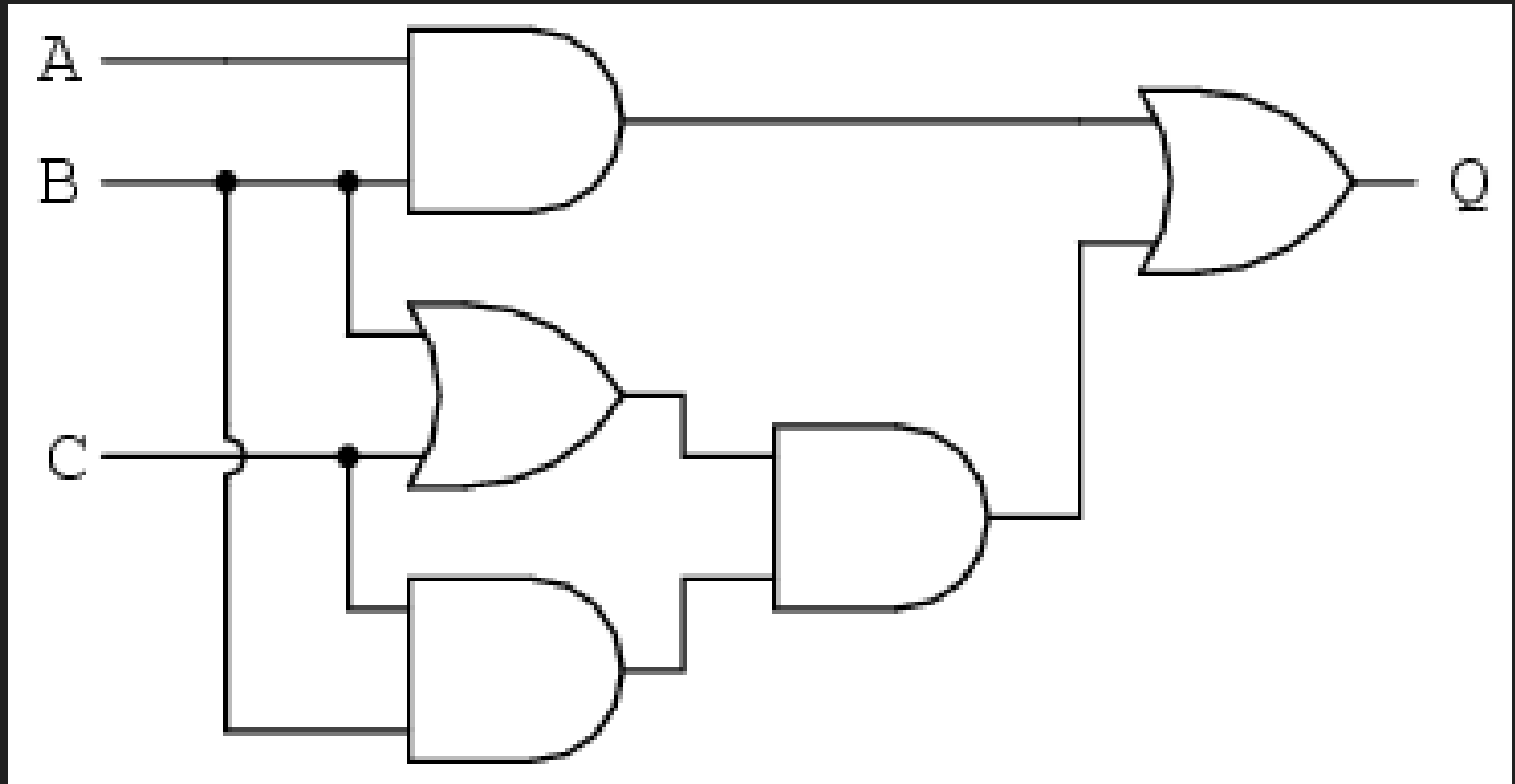
1. Image Signal Processor
2. Video Input (serial, 2 lanes)
3. Safety System
Dual-core lockstep
4. Security System
5. Video Encoder (H.264)
6. GPU (600 GFLOPS, 36 MB SRAM)
7. Processor (12 Cortex-A72)
8. Neural Network Accelerator
96x96 mul/add array
32 MB SRAM
36 TOPS
2 Instances (72 TOPS)
9. LPDDR4 Memory Controller
68 GB/s peak bandwidth
10. Network-on-Chip

Tesla FSD Computer (1000 units)
15W consumed by NNA
72W total to run autopilot
2300 processed frames

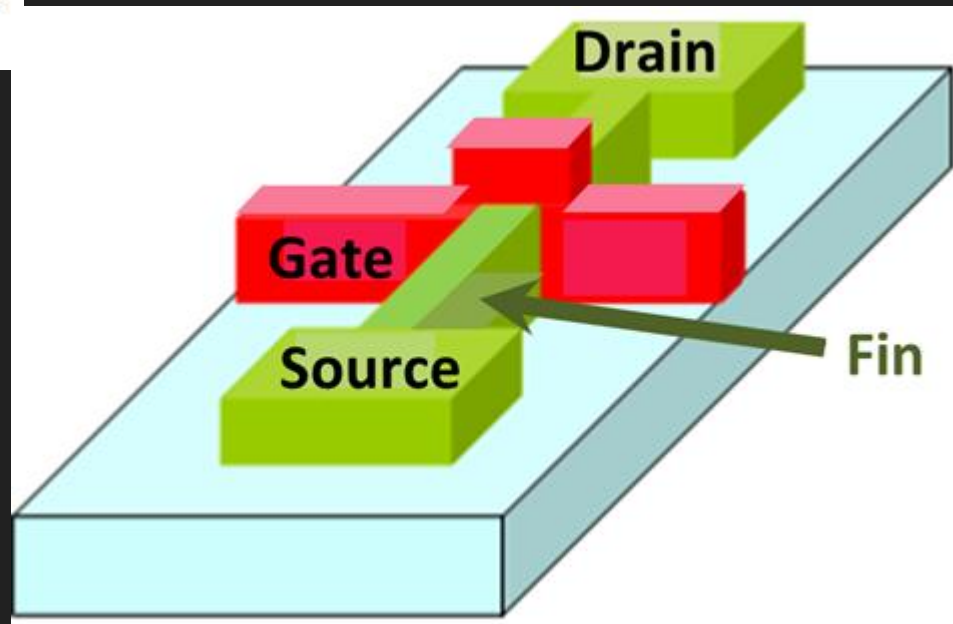
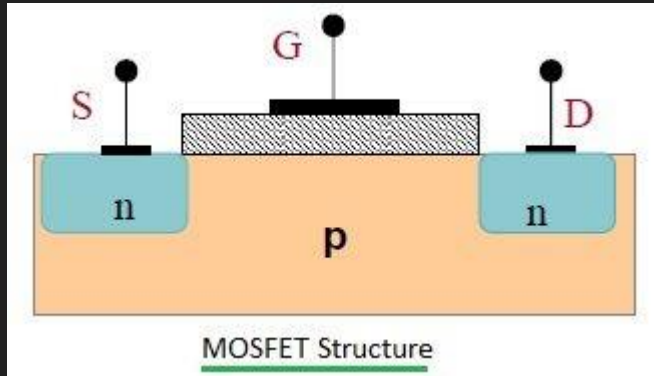
그래서 하는 일은?

"칩"을 만드는 일

하드웨어 개발



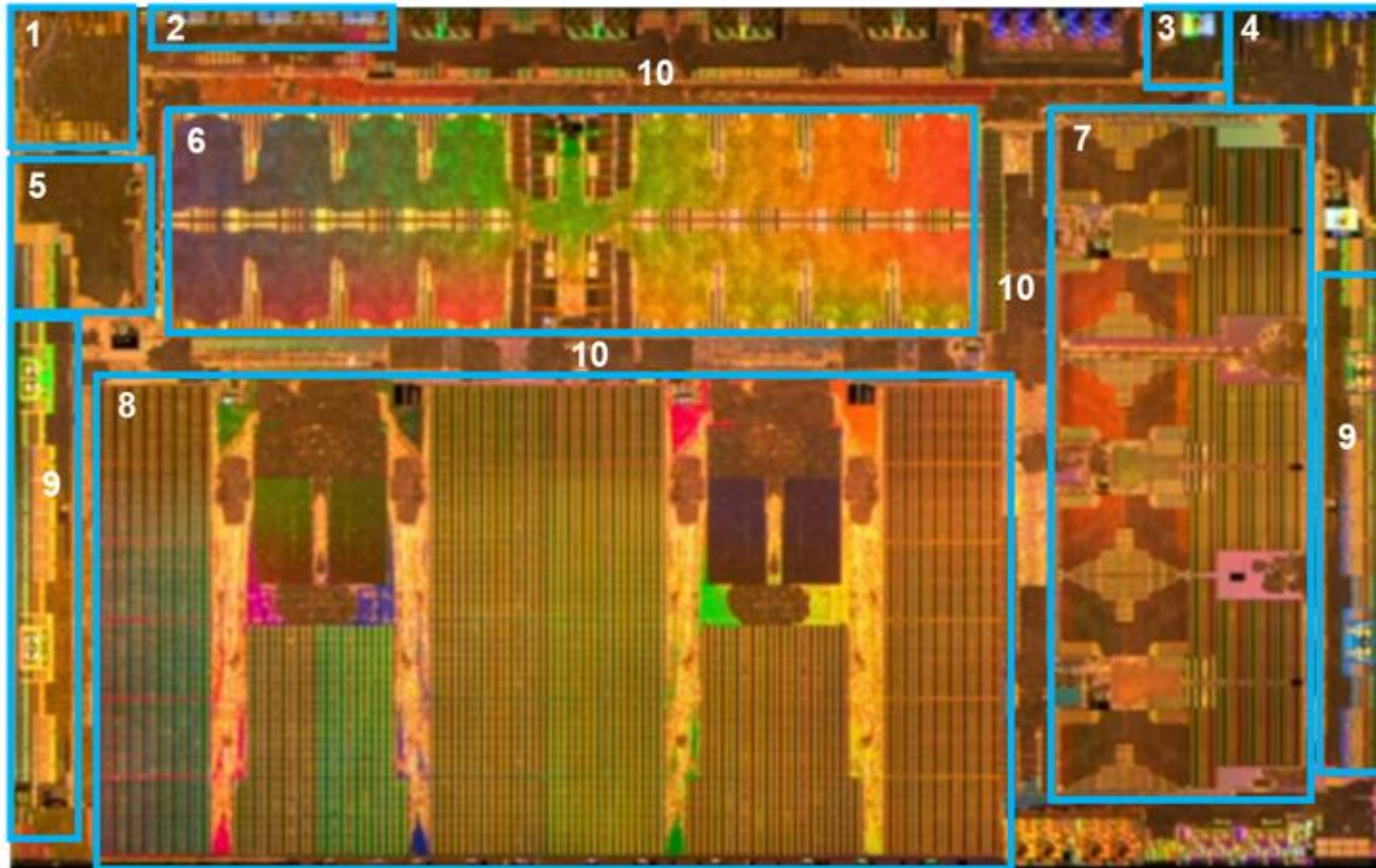
하드웨어 개발



자율주행 (Tesla 칩)

Tesla Full Self-Driving (FSD) Chip

14nm FinFET CMOS (Samsung) – 6 Billion Transistors – 260 mm²



1. Image Signal Processor
2. Video Input (serial, 2 lanes)
3. Safety System
Dual-core lockstep
4. Security System
5. Video Encoder (H.264)
6. GPU (600 GFLOPS, 36 MB SRAM)
7. Processor (12 Cortex-A72)
8. Neural Network Accelerator
96x96 mul/add array
32 MB SRAM
36 TOPS
2 Instances (72 TOPS)
9. LPDDR4 Memory Controller
68 GB/s peak bandwidth
10. Network-on-Chip

Tesla FSD Computer (1000 units)
15W consumed by NNA
72W total to run autopilot
2300 processed frames

거기서
게임 개발하던 프로그래머가
무슨 일을 하나요?

칩을 만들어 보는 비용이 비쌘

게이트의 수 = 수 억 ~수십 억

NVIDIA TU102 = 186억 게이트

한번 시도하는 것에 3개월 이상, 수십 억원 소요

"실물" 칩을 만드는 과정이므로
잘못 만들면 동작하지 않을 수 있다.

모든걸 다 미리 계산, 시뮬레이션, 검증

단위 테스트, 통합 테스트 등의 기능 테스트부터

칩 크기 예측

전력 소모량 시뮬레이션

발열량 시뮬레이션

물리적 특성 검증

심지어 양자 터널링까지..?

시뮬레이션 속도 <<< 실제 속도

완성된 칩에서 5ms 정도 시간이 걸리는 모델
목표하는 칩 클럭이 1~2GHz

현재 CPU를 통한 시뮬레이션 = 약 15시간

약 10,000,000 (=천만)배 느림

어차피 컴퓨팅 파워는 남아도니까..

"훨씬" 빠른 시뮬레이터 구현 중

클라우드 등을 써서라도

Verilog → C++ 모듈 → 테스트 코드에서 호출

기존 툴은 컴파일에만 메모리 100GB 이상 사용

* 유료 시스템도 있음

10억개 정도의 게이트를 가진 회로를

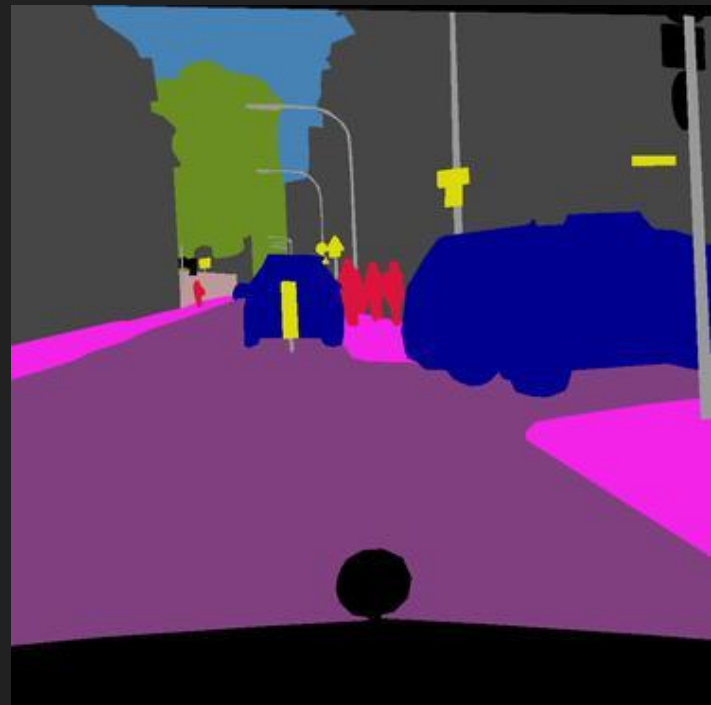
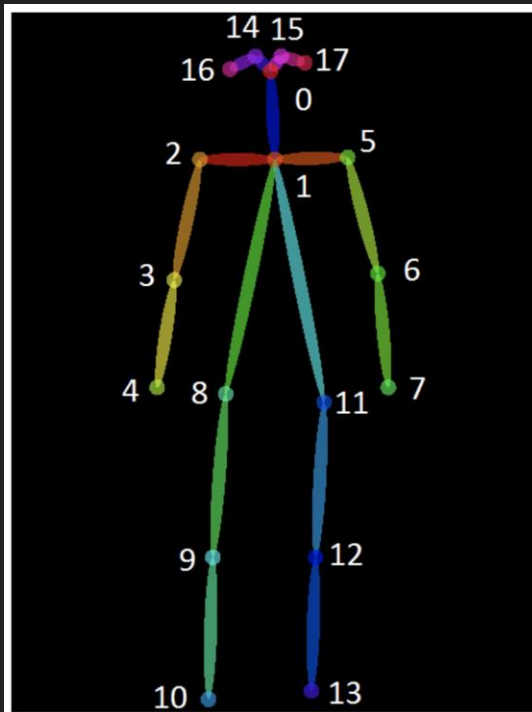
실제 칩 대비 1000~10,000 배 느리게 시뮬레이션

매우 비쌌. 대당 약 100억

데모 개발

~ 특이한 기능을 쓸 수 있는 게임 프로토타이핑

HTML5 canvas, Unity, ...



끝.