



## واحد پردازش عصبی چیست و چه کاربردی در تلفن همراه دارد؟

کدپستی ۱۵۸۳۶۱۶۴۱۴  
[www.rahaco.net](http://www.rahaco.net)

آدرس: تهران، خیابان سپهبد قرنی، خیابان دهقانی، پلاک ۱۲  
تلفن: ۰۲۱۵۴۵۲۱

## فهرست

3	..... واحد پردازش عصبی چیست؟
3	..... اولین چیپست واحد پردازش عصبی.
3	..... استفاده از واحد پردازش عصبی در محصولات میکروسافت
4	..... قابلیت‌های واحد پردازش عصبی NPU
4	..... چرا به واحد پردازش عصبی NPU نیاز است؟
4	..... NPU برای چه اهدافی مورد نیاز است؟
5	..... تفاوت GPU و NPU
5	نتیجه گیری

در حال حاضر که هوش مصنوعی به شکل پیوسته پر رنگ می‌شود. وجود یک پردازنده اختصاصی جهت عملیات پردازش سنگین در این حوزه، درون گوشی‌ها می‌تواند شکل تعامل ما با این گجت را تغییر دهد. اکثر تولیدکننده‌های چیپست موبایل دیر یا زود بحث پردازش هوش مصنوعی را در تولیدات خود به شکل ویژه قرار خواهند داد. هوآوی و اپل از اولین شرکت‌هایی هستند که این فناوری را در چیپست‌های خود به کار بستند. چیپست کایرین 970 و اپل A11 با واحد پردازش عصبی یا همان NPU عرضه شده‌اند اما سامسونگ و کوالکام هنوز چنین پردازنده‌ای ندارند.

## واحد پردازش عصبی چیست؟

واحد پردازش عصبی (NPU) یک ریزپردازنده است که در شتاب‌دهنده‌ها به الگوریتم‌های یادگیری ماشین تخصص دارد که به عنوان پردازنده عصبی شناخته می‌شود. توجه به این نکته ضروری است که نمی‌توان از آن برای محاسبات عمومی مانند واحد پردازش مرکزی (CPU) استفاده کرد. که عمدتاً به این دلیل می‌باشد که پشتیبانی نرم افزاری برای این دسته از پردازنده‌ها توسعه نیافته است که می‌تواند برای هر هدف محاسباتی مورد استفاده قرار گیرد. در واقع توسعه چنین نرم افزار/کامپایلری می‌تواند یک چالش باشد. و در عین حال ممکن است برای کارهایی که برای آن‌ها طراحی نشده است عملکرد پایینی داشته باشد.

## اولین چیپست واحد پردازش عصبی

اولین چیپست با واحد پردازشگر عصبی اختصاصی NPU است. تراشه هوش مصنوعی کایرین در سال 2017 رونمایی شد. این پردازنده یک چیپست متفاوت از سایر پردازنده‌ها می‌باشد. توسعه برنامه‌های هوش مصنوعی بر پایه چیپست کایرین 970 بر روی گوشی‌ها، نیازمند توسعه دادن منابع مورد نیاز نرم‌افزاری در گوشی‌های هواوی است. واحد پردازشی کایرین 970، یک هوش مصنوعی بر اساس فناوری Cloud را در کنار هوش مصنوعی بر روی گوشی‌ها ارائه کرده است.

این سیستم آنالیز داده که در کنار CPU و GPU جانمایی شده است تا 25 برابر کارایی بالاتری را در پردازش‌های گرافیکی و منطقی در گوشی ارائه می‌کند. نقطه قوت این موضوع اینجاست که این افزایش عملکرد با یک افزایش بهره‌وری همراه شده است. به عبارتی دیگر در کایرین 970 هر اندازه که کارایی بیشتر شده باشد، مصرف انرژی آن به نصف کاهش یافته است. این گونه که مشاهده می‌شود قرار دادن NPU در این چیپ، سبب افزایش کارایی شگفت‌انگیزی در بخش‌های پردازش مرکزی و گرافیکی آن شده است.

## استفاده از واحد پردازش عصبی در محصولات مایکروسافت

اضافه شدن یک NPU به تمام دستگاه‌های نسل جدید سرفیس اعلام شده است. مایکروسافت با اضافه کردن یک واحد NPU به محصولات بعدی قصد دارد از این طریق سرعت پردازش برنامه‌های مرتبط با هوش مصنوعی و

یادگیری ماشینی را در این محصولات افزایش دهد. گزارش‌های پیش از این نیز نشان می‌دهد این شرکت پس از راه اندازی بینگ چت تمرکز گسترده‌ای را روی هوش مصنوعی گذاشته است. گزارش جدید نشان می‌دهد رایانه‌های شخصی سرفیس در آینده مجهز به پردازنده‌های اینتل، AMD و تراشه SQ3 مبتنی بر آرم خواهند بود. البته تراشه مایکروسافت از روی چیپ اسنپدراگون ۸ نسل ۳ توسعه می‌یابد.

## قابلیت‌های واحد پردازش عصبی NPU

گوشی‌های مجهز به فناوری NPU می‌توانند عادت‌های کاربر از تلفن همراه خود را در هنگام استفاده شناسایی کرده و مصرف انرژی خود را مطابق با آن تنظیم کند. به عنوان مثال: اگر گوشی شما بفهمد که شما معمولاً ساعت ده شب گوشی را برای شارژ کردن به پریز برق متصل می‌کنید، مصرف انرژی را به نحوی مدیریت می‌کند که تا ساعت ده شب دوام بیاورد.

واحد پردازش عصبی (NPU) عموماً در گوشی‌های پیشرفته و گران قیمت استفاده می‌شود، وظیفه‌ای که یک پردازشگر شبکه عصبی بر عهده دارد به حوزه هوش مصنوعی ارتباط دارد مانند: تشخیص دادن صوت، تصویر و جداسازی سوژه‌ها که این موارد در عکاسی کاربردی و مهم است. روی تراشه‌های گوشی معمولاً تعدادی عدد با واحد نانومتر درج شده است، این اعداد نمایانگر ابعاد یک تراشه است و هر اندازه که کوچک‌تر باشد یعنی آن تراشه سیم کشی داخلی کمتری دارد. این موضوع به همراه فناوری NPU بهینه بودن تراشه و مصرف پایین انرژی آن تراشه را نشان می‌دهد.

## چرا به واحد پردازش عصبی NPU نیاز است؟

در CPU و GPU پردازش اطلاعات به شکل گام به گام انجام می‌شود. یعنی پس از انجام مرحله الف به سراغ مرحله ب می‌رویم. اما در مقابل عملکرد NPU مانند مغز انسان است و همه چیز را همزمان و در لحظه پردازش می‌کند. هدف اصلی از ساخت NPU پردازش محلی عملیات هوش مصنوعی است. طی سال‌های گذشته پیشرفت باور نکردنی در برنامه‌های یادگیری ماشینی به ثمر رسیده و در برخی از موارد مانند بازی‌ها از انسان‌ها بهتر عمل کرده است. در عین حال، برنامه‌های یادگیری ماشینی زندگی انسان را به سطح بالاتری می‌برد.

## NPU برای چه اهدافی مورد نیاز است؟

NPU در مقایسه با پردازنده‌های گرافیکی، محاسبه وظایف یادگیری ماشینی را چندین برابر (نزدیک به 10 هزار بار) سریع‌تر انجام می‌دهد. صرفه جویی در مصرف انرژی و بهبود استفاده از منابع برای وظایف یادگیری ماشینی در مقایسه با GPU و CPU به عنوان مثال پیاده سازی واقعی واحدهای پردازش عصبی (NPU) عبارتند از:

- TPU توسط گوگل

- NNP، Myriad، EyeQ توسط اینتل
- NVDLA توسط Nvidia
- AWS Inferentia توسط آمازون
- Ali-NPU توسط Alibaba
- Kunlun توسط Baidu
- IPU توسط Graphcore
- موتور عصبی توسط اپل

## تفاوت GPU و NPU

اگر چه GPU دارای مزیت در قابلیت محاسبات موازی است، اما به تنهایی کار نمی‌کند و نیاز به پردازش مشترک CPU دارد. ساخت مدل‌های شبکه عصبی و جریان‌های داده هنوز بر روی CPU انجام می‌شود. هرچه عملکرد بالاتر، GPU بزرگ‌تر، مصرف برق بیشتر و گرانتر می‌باشد که برای برخی از دستگاه‌های کوچک و تلفن‌های همراه قابل استفاده نخواهد بود. در مقابل NPU دارای اندازه کوچک، مصرف انرژی کمتر، عملکرد محاسباتی و راندمان بالاتر است.

NPU با شبیه سازی نورون‌ها و سیناپس‌های انسانی در لایه مدار و پردازش مستقیم نورون‌ها و سیناپس‌ها در مقیاس بزرگ با مجموعه دستورالعمل‌های یادگیری عمیق، که در آن یک دستورالعمل پردازش مجموعه‌ای از نورون‌ها را تکمیل می‌کند، کار می‌کند. در [مقایسه با CPU](#) و GPU، NPU ذخیره سازی و محاسبات را از طریق وزن‌های سیناپسی یکپارچه می‌کنند. بنابراین کارایی عملیاتی را بهبود می‌بخشند. پردازنده‌های CPU و GPU باید از هزاران دستورالعمل برای تکمیل پردازش نورون استفاده کنند NPU. را می‌توان تنها با یک یا چند دستورالعمل تکمیل کرد، بنابراین مزایای آشکاری در کارایی پردازش یادگیری عمیق دارد.

## نتیجه گیری

به طور خلاصه و مفید این تراشه‌ها وظایف مربوط به یادگیری ماشینی یادگیری عمیق و شبکه‌های عصبی را با دقت و سرعت بیشتری نسبت به CPU و GPU انجام می‌دهند. با استفاده از واحد پردازش عصبی، موبایل‌ها و دیگر گجت‌های دنیای تکنولوژی رابطه بهتری با اعداد و ارقام پیدا کرده و فناوری‌های سودمندی همچون بهبود کیفیت تصویر، جستجوی محتوای ویدیوها، تشخیص اشیا و مانند این‌ها را با دقت و سرعت بیشتری انجام خواهند داد.

مجله  
رهاکو



رهاکو، مرجع تخصصی مجازی سازی ایران

# مجله رهاکو

RAHA MAG

آدرس: تهران، خیابان سپهد قرنی، خیابان دهقانی، پلاک 12  
کدپستی 1583616414      تلفن: 02154521      [www.rahaco.net](http://www.rahaco.net)

