**آموزش شبکه‌های عصبی جهت پردازش تصویر**

**مقدمه**

در این فصل از پایان‌نامه، ابتدا به تاریخچة مختصری درباره الگوریتم‌های پردازش تصویر اشاره شده است. سپس درباره سیر تکامل شبکه‌های عصبی کلاس‌بندی و تشخیص اشیا پرداخته شده است.

سپس در ادامة فصل مراحل آموزش شبکه عصبی یولو توضیح داده شده است. در آخر به اعتبارسنجی و بررسی سرعت مدل آموزش‌دیده پرداخته شده است.

**الگوریتم‌های پردازش تصویر**

**الگوریتم‌های کلاسیک**

پیش از استفاده از شبکه‌های عصبی در پردازش تصویر، الگوریتم‌های کلاسیک استفاده می‌شدند. این الگوریتم‌ها بر پایة عملیات ریاضی مانند تبدیل فوریه هستند. الگوریتم‌های کلاسیک به‌تنهایی در تشخیص اشیا بسیار محدود هستند. امروزه از الگوریتم‌های کلاسیک بیشتر برای کاهش نویز و پیش‌پردازش دادگان استفاده می‌شود. نمونه‌ای از الگوریتم‌های کلاسیک تشخیص لبه با استفاده از تبدیل فوریه سریع در شکل 1 قابل‌مشاهده است.

(شکل فوریه)

**الگوریتم‌های بر پایة شبکه عصبی**

با پیشرفت شبکه‌های عصبی و قوی‌تر شدن کامپیوترها و معرفی شبکه‌های عصبی کانوولوشنی (شکل 1)، شرایط جایگزینی شبکه‌های عصبی با الگوریتم‌های کلاسیک فراهم شد. شبکه‌های عصبی کانوولوشنی باهدف کلاس‌بندی تصاویر مورداستفاده گرفتند. در ادامه این شبکه‌ها توسعه پیدا کردند و به شکل شبکه‌های عصبی کانوولوشنی مبتنی بر منطقه در آمدند که قابلیت تشخیص اشیا داشت.

(تصویر CNN)

مشکل اصلی شبکه‌های کانوولوشنی مبتنی بر منطقه سرعت پایین و حجم بالای پردازشی آن‌ها بود. شبکه‌های عصبی کانوولوشنی مبتنی بر منطقه سریع و شبکه‌های عصبی کانوولوشنی مبتنی بر منطقه سریع‌تر (شکل 1) باهدف بهبود سرعت و دقت این شبکه‌ها ابداع شدند.

علت اصلی سرعت پایین این نوع شبکه‌ها این بود که الگوریتم آن‌ها بر پایة پردازش شبکه عصبی کانوولوشنی در حدود 2000 بار در نواحی مختلف یک تصویر بود.

(تصویر rcnn)

**شبکه‌های عصبی یولو**

شبکه‌های عصبی یولو (You Only Look Once) یکی از الگوریتم‌های پردازش تصویر بر پایه شبکه‌های عصبی هستند که باهدف بهبود سرعت و دقت در تشخیص اشیا ابداع شدند.

(تصویر یولو 1)

یولو، در مقابل شبکه‌های عصبی کانوولوشنی مبتنی بر منطقه، فقط یک‌بار به تصویر نگاه می‌کند. این کار باعث می‌شود که سرعت پردازش تصاویر بسیار بالاتر برسد و حجم پردازشی کاهش یابد. یولو تصویر را به یک شبکه کانوولوشنی می‌فرستد و در نهایت یک بردار ویژگی را برمی‌گرداند. این بردار ویژگی شامل اطلاعات مربوط به مکان و دسته‌بندی اشیا در تصویر است.

یولو نسبت به شبکه‌های عصبی کانوولوشنی مبتنی بر منطقه، دقت کمتری دارد؛ اما سرعت بسیار بالاتری را ارائه می‌دهد. این ویژگی یولو را برای کاربردهایی که نیاز به پردازش تصویر در زمان واقعی دارند، مناسب می‌کند. از زمان معرفی یولو 1 در سال 2015، هر سال نسخه‌های جدیدتر این مدل معرفی شده‌اند. در این پروژه مدل یولو 8 برای آموزش انتخاب شده است.

**آموزش شبکه عصبی یولو**

**یادگیری انتقالی در یولو**

آموزش شبکه‌های عصبی می‌تواند در بسیاری از مواقع وقت‌گیر و پرهزینه باشد. به همین دلیل یادگیری انتقالی باهدف استفاده از اطلاعات قبلی شبکه عصبی‌های آموزش‌دیده قبلی ابداع شده است. در این روش شبکه عصبی آموزش‌دیده بر مجموعه‌دادگان بزرگ را برداشته و لایه‌های آخر آن را حذف یا تغییر می‌دهند تا در مجموعه‌دادگان جدید آموزش ببیند (شکل 1). یادگیری انتقالی باعث می شود شبکه ی جدید بسیار سریع تر آموزش ببیند که باعث صرفه جویی در هزینه و زمان آموزش شبکه می شود.

(تصویر یادگیری انتقالی: یادگیری انتقالی)

**فاز اول آموزش یولو**

مجموعه‌دادگان ۱۲۰۰تایی که در فاز اول فصل قبلی معرفی شده است برای آموزش و یادگیری انتقالی مدل از پیش آموزش‌دیدة یولو 7 تاینی قرار گرفته است. برای اینکار تغییراتی در مدل از پیش آموزش دیده ی یولو انجام شده است. خروجی مدل آموزش دیده یولو شامل 80 کلاس می باشد. در ابتدا تعداد خروجی ها به 6 کلاس تغییر یافت که به شکل زیر می باشد.

مدل آموزش‌دیده یولو: coco.yaml

مدل تشخیص میوه: fruit-6.yaml

نتایج آموزش نشان داد که این نوع کلاس‌بندی حاصل تمام وزن‌های قبلی را از بین می‌برد. برای مثال، در مدل آموزش‌دیده، کلاس شماره 50 برای تشخیص پرتقال آموزش دیده است ولی هنگام فاین تیون کردن، کلاس سوم برای پرتقال تازه قرار گرفته است که قبلا برای تشخیص ماشین آموزش دیده است. به همین دلیل روش بهتر استفاده از 83 کلاس می باشد تا بتوان از وزن های قبلی هم استفاده کرد که به صورت زیر میباشد.

مدل تشخیص میوه: fruit-83.yaml

در این نوع کلاس‌بندی، کلاس‌های قبلی میوه‌ها به میوه‌های تازه تغییر نام پیدا کردند و 3 کلاس جدید برای تشخیص میوه های فاسد استفاده شده است. نتایج آموزش در شکل 1 قابل مشاهده هستند.

(تصاویر آموزش فاز اول)

نتایج بر روی دادگان اعتبارسنجی مجموعه‌دادگان ۱۲۰۰تایی بالای 95 درصد بود ولی مشکل اصلی عملکرد آن بر روی دیگر دادگانیست که برچسب نداشتند بود. نمونه ای از نتایج بر روی دادگان در شکل 1 قابل مشاهده است.

(عکس مثال‌های اشتباه: نمونه موارد تشخیص اشتباه میوه‌ها)

دو دلیل اصلی باعث به‌وجودآمدن این مشکل می‌شدند. دلیل اول، بسیار متفاوت‌بودن مجموعه‌های دادگان است که نیازمند برچسب‌زدن مجموعه‌های دادگان دیگر بود. دلیل دوم، وجود عکس‌هایی با تعداد زیادی میوه که باعث افزایش مثبت اشتباه (false positive) می‌شوند. یک نمونه از این عکس‌ها در شکل 1 این صورت می باشد.

(تصویر کلی سیب: نمونه تصویر با تعداد بالای میوه‌ها)

به همین سبب، عکس‌هایی با بیش از 5 میوه حذف شدند و مجموعه دادگان 1200 تایی به 1096 عکس کاهش پیدا کرد.

**فاز دوم آموزش یولو**

باتوجه‌به ارائه‌شدن مدل هوش مصنوعی یولو 8 در زمان فاز دوم و عملکرد بهتر این شبکه نسبت به یولو 7، از این شبکه برای آموزش مجموعه دادگان جدید استفاده شده است. نتایج آموزش در شکل 1 قابل رویت هستند.

(تصاویر آموزش فاز دوم)

نتایج نشان می‌دهند که دقت بر روی دادگان اعتبارسنجی بین 95 تا 98 درصد هست. همچنین عملکرد شبکه بر روی عکس‌هایی که قبلاً ندیده است و مدل قبلی در تشخیص آن‌ها ضعف داشته است، درست است.

**راستی‌آزمایی و سرعت سنجی یولو**

**بررسی دقت یولو**

**بررسی سرعت یولو**

**جمع‌بندی**

|  |  |
| --- | --- |
| مدل آموزش‌دیده در فاز اول | مدل آموزش‌دیده در فاز دوم |
|  |  |
|  | A close-up of a green apple  Description automatically generated |
|  | A close-up of a rotten orange  Description automatically generated |
|  |  |