

دو مقاله، LeCun et al. و Zeiler et al.، چندین جنبه مشترک کلیدی در بحث‌های خود درباره شبکه‌های عصبی و تشخیص الگو به اشتراک می‌گذارند:

۱- شبکه‌های عصبی پیچشی (CNNs)

هر دو مقاله به شدت به شبکه‌های عصبی پیچشی^۱ می‌پردازند. لکان و همکاران اصول طراحی آن‌ها را، مانند میدان‌های پذیرش محلی (local receptive fields)، وزن‌های مشترک (shared weights) یا هم‌تاسازی وزن (weight replication)، و زیرنمونه‌گیری مکانی یا زمانی (spatial or temporal sub-sampling) برای دستیابی به تغییرناپذیری نسبت به جابجایی، مقیاس و اعوجاج در وظایفی مانند تشخیص کاراکتر، با جزئیات شرح می‌دهد.

زیلر و همکارش بر اساس این CNN‌های بنیادی، به ویژه مدل‌های بزرگ، وظایف چالش‌برانگیز طبقه‌بندی بصری مانند ImageNet تبیین می‌کنند و عملکرد چشمگیر آن‌ها را تأیید می‌کند.

۲- یادگیری مبتنی بر گرادیان^۲ و پس‌انتشار^۳

یک اشتراک اساسی در این دو مقاله، اتکا به الگوریتم‌های یادگیری مبتنی بر گرادیان است، با پس‌انتشار به عنوان روش اصلی برای آموزش شبکه‌های عصبی است. لکان و همکاران تأکید می‌کنند که شبکه‌های عصبی چندلایه آموزش‌دیده با پس‌انتشار، بهترین نمونه از تکنیک‌های موفق یادگیری مبتنی بر گرادیان هستند. زیلر و همکارش بیان می‌کند که پارامترهای شبکه با پس‌انتشار مشتق تابع زیان (loss function) و به‌روزرسانی از طریق نزول گرادیان تصادفی (stochastic gradient descent) آموزش می‌بینند.

۳- یادگیری خودکار^۴ و سیستم‌های مبتنی بر داده^۵

هر دو منبع از یک تغییر پارادایم به سمت اتکای بیشتر بر یادگیری خودکار و کمتر بر اکتشافات طراحی شده دستی (hand-designed heuristics) برای ساخت سیستم‌های تشخیص الگو حمایت می‌کنند. این امر با حرکت از استخراج ویژگی‌های دست‌ساز (hand-crafted feature extraction) به یادگیری مستقیم ویژگی‌ها از تصاویر پیکسلی، و همچنین جایگزینی ماژول‌های یکپارچه دستی با سیستم‌های یکپارچه و قابل

^۱ Convolutional Neural Networks

^۲ Gradient-Based Learning

^۳ Backpropagation

^۴ Automatic Learning

^۵ Data-Driven Systems

آموزش جهانی مانند Graph Transformer Networks (GTNs) در مقاله لکان و همکاران برجسته شده است. زیلر و همکارش به طور ضمنی با نمایش قدرت ویژگی‌های CNN آموخته‌شده در مجموعه داده‌های پیچیده، از این موضوع حمایت می‌کند.

۴- تغییرناپذیری نسبت به تبدیل‌ها و اعوجاج‌ها^۶

یک جنبه حیاتی برای تشخیص الگوی قوی، تغییرناپذیری نسبت به واریانس‌ها، جابجایی‌ها، مقیاس‌ها و اعوجاج‌های الگوهای ورودی است. لکان و همکاران توضیح می‌دهند که چگونه CNN‌ها به طور خاص برای مقابله با تغییرپذیری اشکال 2D طراحی شده‌اند، و به طور خودکار از طریق اشتراک وزن (weight sharing) به تغییرناپذیری جابجایی (shift invariance) دست می‌یابند و از طریق لایه‌های زیرنمونه‌گیری^۷ حساسیت به جابجایی‌ها و اعوجاج‌ها را کاهش می‌دهند. زیلر و همکارش تغییرناپذیری‌های آموخته‌شده ویژگی‌ها را در لایه‌های بالاتر CNN نشان می‌دهد و حساسیت مدل به مکان شیء به جای صرفاً زمینه کلی صحنه را از طریق آزمایش‌های انسداد (occlusion experiments) تأیید می‌کنند.

Ref:

[1] Yann LeCun Leon Bottou Yoshua Bengio and Patrick Haffner, Gradient Based Learning Applied to Document Recognition, PROC OF THE IEEE NOVEMBER 1998.

[2] Matthew D. Zeiler and Rob Fergus, Visualizing and Understanding Convolutional Networks, European Conference on Computer Vision 2014 Sep 6 (pp. 818-833). Cham: Springer International Publishing.

^۶ Invariance to Transformations and Distortions

^۷ sub-sampling layers