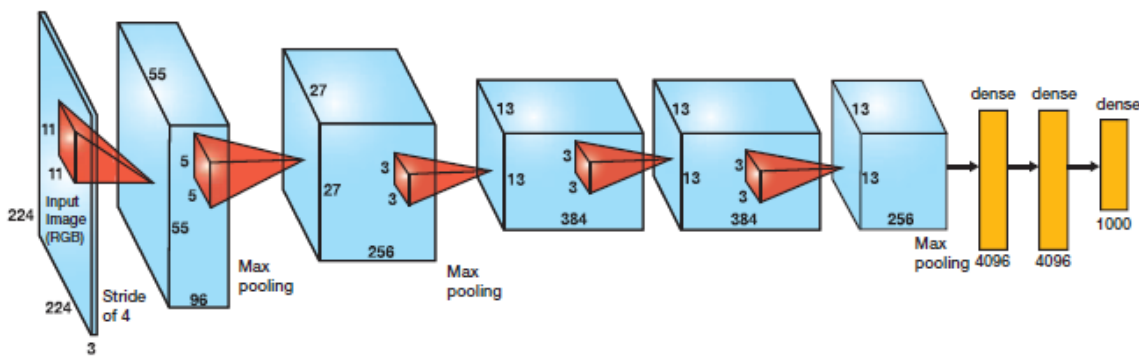


AlexNet

AlexNet เป็น deep learning model ตัวคลาสสิก ที่ถูกสร้างโดย Geoffrey Hinton ใช้สถาปัตยกรรมแบบง่าย ๆ โดยมี convolutional layer และ pooling layer ประกอบกัน และมี fully connected layer อยู่บนสุด ปัจจุบัน ยังใช้เป็น โมเดลต้นแบบ เป็นตัวตั้งต้นของการทำงานคอมพิวเตอร์วิชั่น

AlexNet เป็นหนึ่งในเครือข่ายลึกแรกที่จะผลักดันความถูกต้องของการจัดหมวดหมู่ ImageNet โดยก้าวสำคัญในการเปรียบเทียบกับวิธีการแบบดั้งเดิม ประกอบด้วย 5 ชั้น convolutional ตามด้วย 3 ชั้นเชื่อมต่ออย่างเต็มที่ตามที่ depicted ในรูป



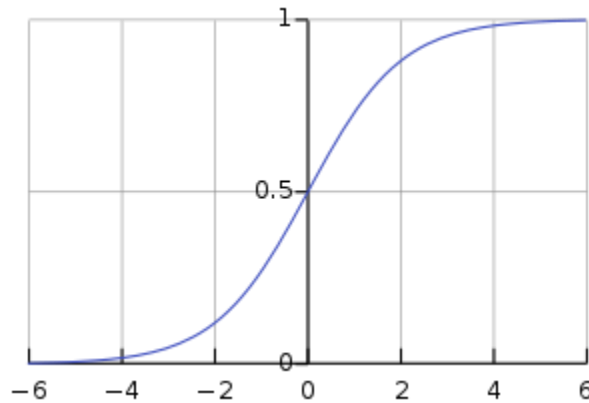
- เลเยอร์ Convolutional ที่ 1: เลเยอร์นี้กรองภาพอินพุต ($224 * 224 * 3$) ด้วย 96 kernels ขนาด $11 * 11 * 3$
- เลเยอร์ Convolutional ที่ 2: เลเยอร์นี้กรองเอาต์พุตของเลเยอร์คอนโวลูชันที่ 1 ซึ่งได้รับการตอบสนองตามปกติ และรวมเข้ากับ 256 เคอร์เนลขนาด $5 * 5 * 96$
- เลเยอร์ Convolutional ที่ 3: ไม่มีการตอบสนองปกติและการรวมกลุ่ม มีเมลิตขนาด $3 * 3 * 256$ จำนวน 384 เมลิต
- เลเยอร์ Convolutional ที่ 4: ไม่มีการทำให้เป็นมาตรฐานการตอบสนองในพื้นที่และการรวมกลุ่ม มี 384 เมลิต ขนาด $3 * 3 * 384$
- เลเยอร์ Convolutional ที่ 5: ในเลเยอร์นี้เอาต์พุตจะรวมกัน มี 256 เมลิตขนาด $3 * 3 * 384$
- ชั้นที่ 1 และ 2 ที่เชื่อมต่อกันอย่างเต็มที่: แต่ละชั้นมีเซลล์ประสาท 4096 เซลล์
- เลเยอร์ที่เชื่อมต่ออย่างสมบูรณ์ที่ 3: นี่คือการเอาต์พุต "ป้อนเข้ากับซอฟต์แวร์แม็กซ์ 1000 ทางซึ่งสร้างการกระจายมากกว่าฉลากคลาส 1000"

AlexNet เสนอโดยอเล็กซ์ Krizhevsky ใช้ ReLu (หน่วยเส้นตรง Rectified) สำหรับส่วนที่ไม่ใช่เชิงเส้น, แทนการแทนแทน ฟังก์ชัน Tanh หรือ Sigmoid ซึ่งเป็นมาตรฐานก่อนหน้านี้สำหรับเครือข่ายประสาทแบบดั้งเดิม. ReLu จะได้รับโดย

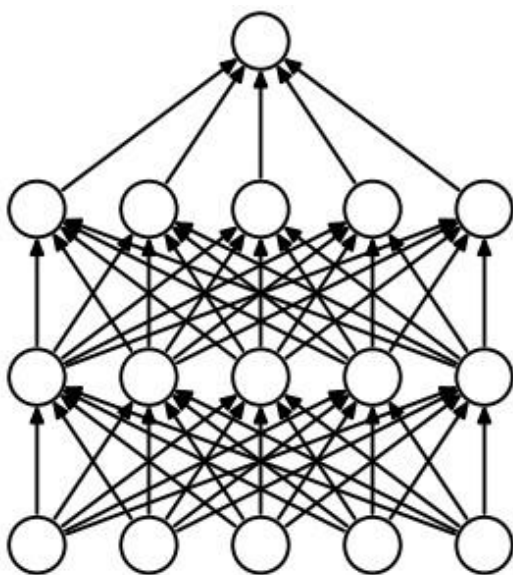
$$f(x) = \text{สูงสุด}(0, x)$$

ข้อดีของ ReLu มากกว่า sigmoid คือรถไฟเร็วกว่าหลังเพราะอนุพันธ์ของ sigmoid กลายเป็นขนาดเล็กมากในภูมิภาคอิมตัวและดังนั้นจึงปรับปรุงน้ำหนักเกือบหายไป นี่เรียกว่าหายไปปัญหาการไล่ระดับสี

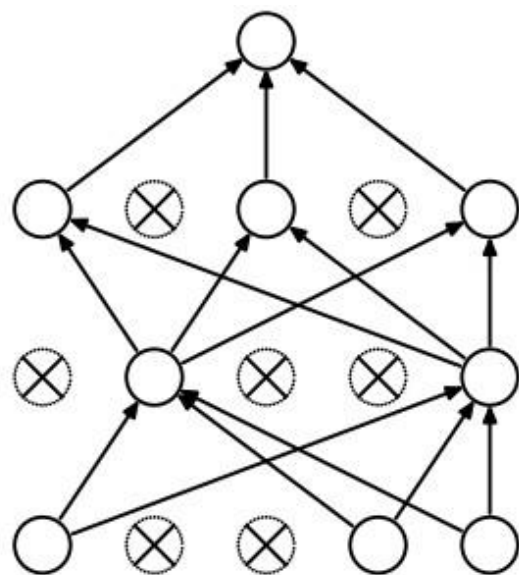
ในเครือข่าย ReLu ชั้นจะถูกใส่หลังจากที่แต่ละและทุกชั้น convolutional และเชื่อมต่ออย่างเต็มที่ (FC)



ปัญหาอื่นที่สถาปัตยกรรมนี้แก้ไขได้ลดเกินเหมาะสมโดยใช้ชั้น Dropout หลังจากทุกชั้น FC เลเยอร์ Dropout มีความน่าจะเป็น(p)ที่เกี่ยวข้องกับมันและถูกนำไปใช้ที่เซลล์ประสาทของแผนที่การตอบสนองที่แยกจากกันทุก มันสุ่มปิดการเปิดใช้งานกับ p น่าจะเป็นที่สามารถมองเห็นในรูป



(a) Standard Neural Net



(b) After applying dropout.

ความคิดที่อยู่เบื้องหลังการเลื่อนจะคล้ายกับรูปแบบวงดนตรี เนื่องจากชั้น dropout, ชุดที่แตกต่างกันของเซลล์ประสาทที่มีการปิด, เป็นตัวแทนของสถาปัตยกรรมที่แตกต่างกันและสถาปัตยกรรมที่แตกต่างกันเหล่านี้ได้รับการฝึกฝนควบคู่กันน้ำหนักให้กับแต่ละชุดย่อยและผลรวมของน้ำหนักเป็นหนึ่ง. สำหรับ n เซลล์ประสาทแนบกับ DropOut จำนวนของสถาปัตยกรรมชุดย่อยที่เกิดขึ้นคือ 2^n ดังนั้นจำนวนการคาดการณ์เฉลี่ยเป็นวงเหล่านี้ของแบบจำลอง นี้ให้ระเบียบ

รูปแบบโครงสร้างซึ่งจะช่วยในการหลีกเลี่ยงการเกิน- กระชับของ มุมมองของ DropOut เป็นประโยชน์ก็คือเนื่องจาก เซลล์ประสาทจะถูกเลือกแบบสุ่ม, พวกเขามักจะหลีกเลี่ยงการพัฒนาการปรับตัวร่วมในหมู่พวกเขาจึงทำให้พวกเขาพัฒนา คุณสมบัติที่มีความหมาย, เป็นอิสระจากคนอื่น ๆ