DAY 13 · Part 2 · How We Got Here

One photo contest flipped the whole game — 2012: deep learning comes back

Five-Minute AI Lessons · Sep 19, 2026

Not a new idea — the day data, computing and depth finally came together

In a nutshell

1. In the 2012 ImageNet photo-recognition contest, a University of Toronto team's neural network won with a **top-5 error of 15.3%**. Second place was 26.2%, a gap of more than 10 percentage points — a first for that contest.

2. Almost nothing was newly invented. **Millions of photos · gaming graphics cards · a deeply stacked neural network** — that year, those three came together for the first time.

3. After that, research on images switched over to neural networks across the board. **It was the day the second winter truly ended.**

When the three came together, the error rate plunged

Data
ImageNet — millions of photos labeled by people

+

Computing
Two gaming graphics cards — fully trained in 5–6 days

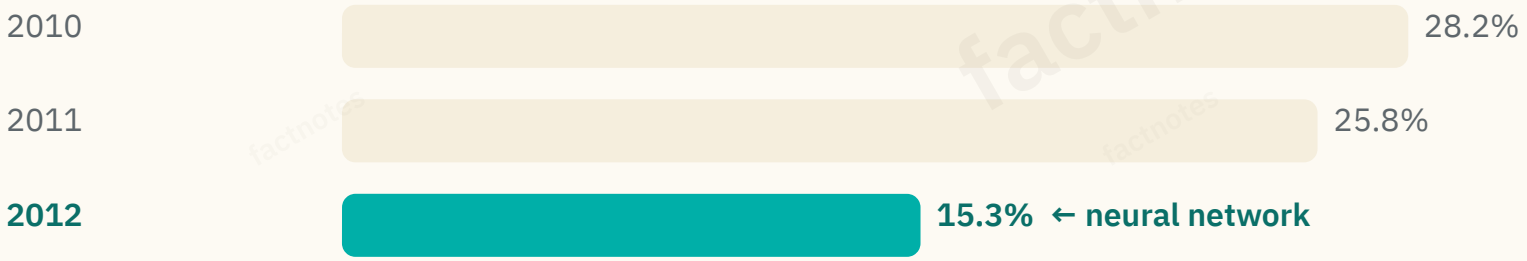
+

Depth
A network eight layers deep — finds features by itself

=

2012 · contest win
15.3% error — beat 2nd place by more than 10 points

ImageNet contest top-5 error rate



Even 2nd place (26.2%) was about the same as before — only one entry plunged.

Key idea: what ended the winter wasn't a new idea — it was three things coming together

What changed on either side of 2012

Scores

Before

It dropped 1–2 points a year.
2010: 28.2% → 2011: 25.8%.
The mood: ‘this is about the limit.’

After

2012: 15.3%.
It dropped 10+ points in one year.
That year's 2nd place: 26.2%.

What was shrinking slowly suddenly plunged.

Finding features

Before

People designed hints like
‘cats have pointy ears’ by hand
and fed them to the computer.

After

Show it tons of photos, and
the network found by itself
which patterns to look at.

Like DAY 08's prompts — not telling it, but letting it find out.

Hardware

Before

On an ordinary CPU,
a network this big took months.
So nobody tried it.

After

On two gaming graphics cards,
done in 5–6 days.
They could try it, so they did.

The part that drew game screens became research gear.

None of the three was a ‘new invention’ — what already existed could finally be run.

Sources: ILSVRC 2012 results (image-net.org) · Krizhevsky, Sutskever, Hinton, ImageNet Classification with Deep CNNs, NIPS 2012

Quick check

Just three questions

Q1 Why was the 2012 ImageNet result such a shock?

① The winner hit 0% error, a first ever
② Winner 15.3% vs. 2nd 26.2%: a 10+ point gap
③ People entered the contest for the first time
④ Only one team entered that year

Q2 What's the biggest reason deep learning only took off in 2012?

① Neural networks were invented that year
② Photos, graphics cards, deep nets finally met
③ The internet was invented that year
④ The government made AI research mandatory

Q3 Why were gaming graphics cards (GPUs) used for deep learning?

① They make screens look nice
② They do tons of math at once
③ The internet got faster
④ They have lots of storage

Know the why, too

Q1

15.3% vs. 26.2%. In a contest that had improved by 1–2 points a year, one team pulled ahead on its own by more than 10 points.

Q2

Neural networks, and the way to stack and fix layers, already existed by 1986 (DAY 12). What was missing was data and computing, and in 2012 both were filled in.

Q3

A graphics card is built to calculate tons of screen dots at the same time. Neural network math has the same shape, so it could be used as-is.

Today in one line

What ended the winter wasn't a new idea —
it was the day data, computing and depth came together.



Photos got easy — but what about language? Tomorrow: the 2017 Transformer. Every AI you use today starts here.