



Chapter 1

• Computer Vision •

AI Vision Deep Dive

How AI sees the world — really!



Welcome to Computer Vision!



Computer Vision = giving AI the gift of SIGHT!

★ What CV Can Do (Class 7 level):

-  • **Classification:** "Is this a cat or dog?"
-  • **Detection:** "Where is the cat?"
-  • **Segmentation:** "Which exact pixels are cat?"
-  • **Generation:** "Make a NEW cat photo!"

Real-world CV products YOU use:



Face Unlock



Photo Tag



Tesla Autopilot



Snapchat Filters



X-ray AI



Traffic Cameras



Amazon Go



AI Art

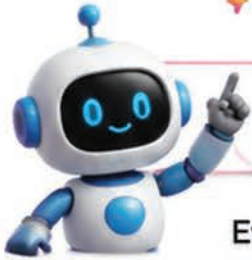


DID YOU KNOW?

Google Photos auto-tags **4 BILLION** photos every day using CV! All without humans!



How AI Sees Pixels!



A photo to AI = numbers in a grid!

Every digital image is made of **PIXELS** – tiny coloured dots!

Each pixel = 3 numbers (RGB):

- Red (0-255)
- Green (0-255)
- Blue (0-255)



Examples:

Colour	RGB	Result
● Pure red	(255, 0, 0)	●
● Pure green	(0, 255, 0)	●
● Pure blue	(0, 0, 255)	●
○ White	(255, 255, 255)	□
● Black	(0, 0, 0)	■
● Indian saffron	(255, 153, 51)	■



A small 100×100 photo:

$100 \times 100 \times 3 = \mathbf{30,000 \text{ numbers!}}$

A full HD photo (1920×1080) = **6.2 MILLION** numbers!



AI reads **ALL** these numbers **FAST** –
and finds patterns!



Classification vs Detection vs Segmentation

3 CV tasks – increasing difficulty!

1



1. CLASSIFICATION

"What is in this picture?"

Output: "Dog!"

2



2. DETECTION

"WHERE is the dog?"

Output: Bounding box
around dog!

3



3. SEGMENTATION

"Which EXACT pixels are dog?"

Output: Pixel-by-pixel mask!



Real-world examples:

Task	Used In	Difficulty
 Classification	Photo album tagging	★ ★
 Detection	Security cameras	★ ★ ★
 Segmentation	Self-driving cars, medical	★ ★ ★ ★ ★

SELF-DRIVING CARS



Tesla uses **ALL THREE** – classifies objects (car/person), detects (where), AND segments (lane boundaries pixel-precise) – **60 times per second!**



Image Segmentation Deep Dive

Segmentation = AI's most powerful vision skill!



Big Word: **SEGMENTATION**

Drawing exact boundaries around objects! Every pixel labeled!

Types of segmentation:



SEMANTIC

Group by class (all cars one colour)



INSTANCE

Each separate object (car 1, car 2, car 3)

Real uses:



Medical: tumour boundaries on MRI scans



Self-driving: road, lane, vehicles, pedestrians



Satellites: forest, water, urban areas



Movies: cut out actors from background!



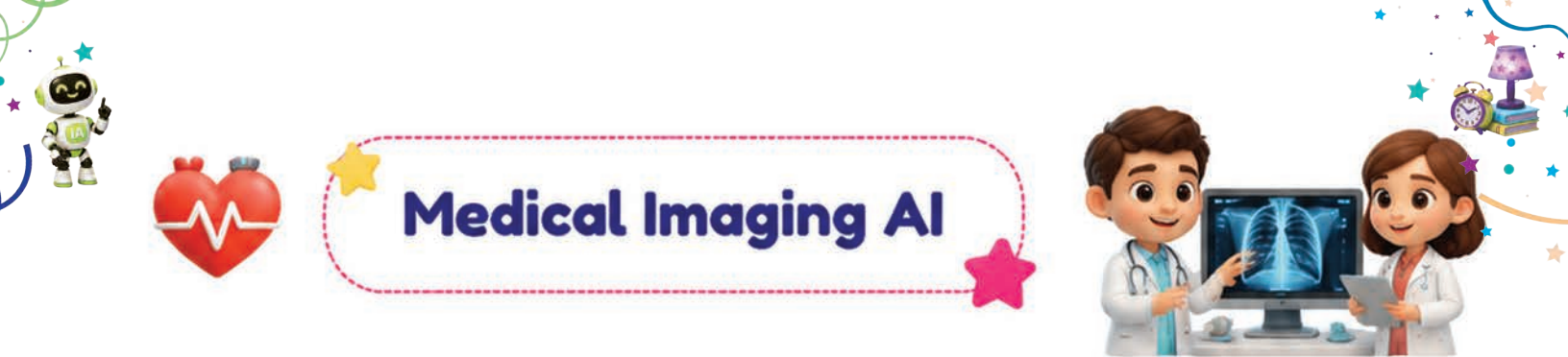
Phone portrait mode: blur background, keep face!



INDIAN AI WIN

AIIMS Delhi uses AI segmentation for brain tumour detection.

95% accuracy! Faster than human radiologists!



Medical Imaging AI

AI is saving lives in hospitals worldwide!

What medical AI can do:



Lung X-rays

Spot TB, pneumonia, cancer



Brain MRI

Tumours, strokes



Eye Scans

Diabetes, glaucoma



Heart Scans

Blockages, arrhythmias



Bone Scans

Fractures



Cell Slides

Cancer cells

How it works: ⚙️

- 1 Train on MILLIONS of labeled scans
- 2 AI learns "normal" vs "abnormal" patterns
- 3 New scan → AI flags risk areas
- 4 Doctor reviews AI's suggestions



Indian success: Qure.ai (Mumbai) AI screens TB in 30 seconds – used in 90+ countries!



AI **ASSISTS** doctors. Final decision is **ALWAYS** the doctor's!

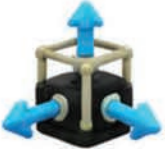


Self-Driving Cars



The ultimate CV challenge – real-time, life-critical!

What a self-driving car SEES (in 1 second):

 8 Cameras 360° view!	 LiDAR/Radar Distance to objects
 GPS + Maps Where am I going?	 IMU sensors Speed & direction

The AI's job (60+ times/sec):

 •SEGMENT roads, lanes, sidewalks	 •DETECT cars, people, signs, lights	 •CLASSIFY all objects	 •PREDICT what they'll do next	 •DECIDE steer/brake/accelerate
--	---	---	--	---

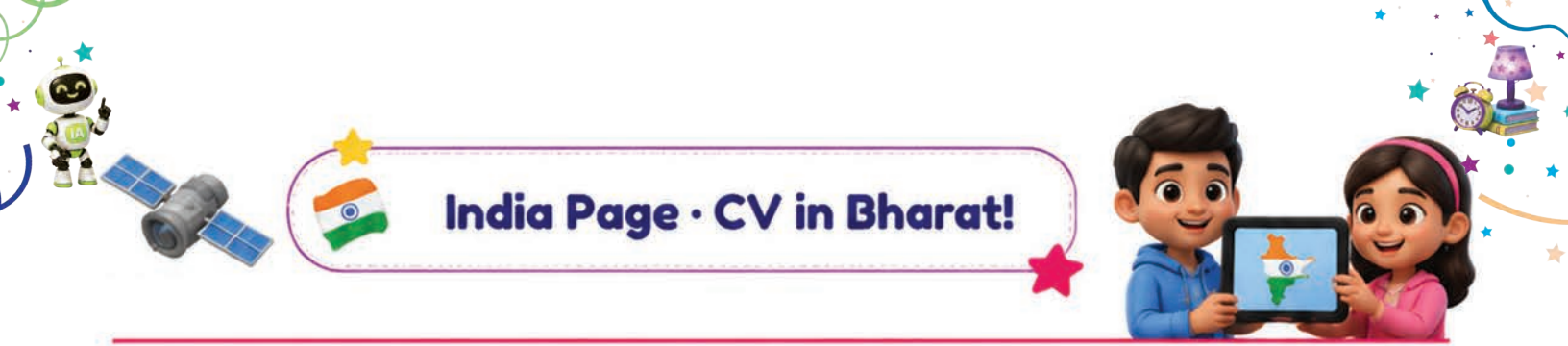
Self-driving levels:

Level	Description	Examples
L0-L1	Driver does all	Old cars, cruise control 
L2	Some AI help	Tesla Autopilot, Honda Sensing 
L3-L4	AI drives mostly	Waymo (US) 
L5	Full autonomy	Future! 



Indian roads are **HARDEST** – animals, traffic, no lane markings! Real challenge!





India Page • CV in Bharat!

Indian companies leading global CV innovation!



Qure.ai

Mumbai • AI for chest X-rays, brain scans • Used in 90+ countries!



Aravind Eye Care

Tamil Nadu • AI detects diabetic retinopathy • 4 million Indians screened!



Bengaluru Smart City

AI cameras count vehicles, manage signals!



KissanAI

Photo of leaf → detects crop disease!



BharatNet AI

Rural identity verification via face



ISRO Bhuvan

Satellite CV – forests, fires, floods!



BHARATNET

India's massive project to verify rural identities & deliver government services via AI face recognition. Reaches 6 LAKH+ villages!



YOUR Idea: Where could CV help your local area?

My CV idea: _____



Generative AI · Creating Images!



AI doesn't just see – it **CREATES**!



Big Word: **GENERATIVE AI**

AI that **CREATES** new content –
images, music, text – that never existed!

Famous image generators:



DALL-E 3

OpenAI · text → image



MidJourney

Artistic styles



Stable Diffusion

Open source, free!



Adobe Firefly

India-friendly



Imagen (Google)

Photo-realistic



Sora

Text → VIDEO!

How it works (simplified):

- 1 Trained on BILLIONS of images + descriptions
- 2 Learns “what cat looks like”, “what red means”
- 3 You type prompt: “cat in space”
- 4 AI creates NEW image matching!



Good prompt example:



Bad prompt “cat picture”

Great prompt! “A fluffy orange cat sitting on a wooden table by a window during sunset, photo-realistic, warm lighting, 4k quality, professional photography”



India's response: 2023 IT Rules require
platforms to label AI content!





Deepfakes & Dangers!



Generative AI's DARK side!



Big Word: DEEPFAKE

Fake videos/images of REAL people doing/saying things they NEVER did!



DANGERS

2024: Deepfakes of Indian celebrities Rashmika Mandanna, Aamir Khan went viral!
Political deepfakes during elections! Real harm done!

Where deepfakes harm:



Elections

Fake political speeches



Fraud

Fake CEO voice calls!



Harassment

Fake images of people



Fake News

Disinformation



How to spot a deepfake:

-  Strange eye movements/blinking
-  Robotic voice tone
-  Weird lighting on face
-  Low resolution mouth area
-  Check source: Trusted news?



India's response: 2023 IT Rules require platforms to label AI content!

Portfolio Project #1 · Build Your First CV Model!



Mission: Train AI to recognize 3 categories using Google's FREE Teachable Machine!

Step-by-step:

1

Go to
teachablemachine.withgoogle.com
Free, in browser, no install!



2

Pick "Image Project" → Standard
Or use Pose/Sound!



3

Create 3 classes
e.g. Apple, Banana, Mango



4

Click "Webcam" – 50 photos each
Show class to camera!



5

Click "Train Model"
Wait 2 min!



6

Test & Export!
Save link/model for portfolio



Portfolio entry:

My 3 classes: _____ Accuracy %: _____

Trained on: _____ images Time taken: _____

Real-world use case for this model: _____



Save screenshot to your portfolio folder!



Ch 1



Build a Face Filter (Concept!)



How Snapchat/Instagram filters work!



3-step recipe:

- 1 **FACE DETECTION:** Find face in camera
- 2 **LANDMARK DETECTION:** Find 68 key points (eyes, nose, lips)
- 3 **OVERLAY:** Place filter on landmarks!

The 68 facial landmarks:



Eyes

Points 37–48



Nose

Points 28–36



Lips

Points 49–68



Eyebrows

Points 18–27



Jaw

Points 1–17



Total

68 unique!



Design YOUR filter:

What filter would you make? Sunglasses on eyes?
Dog ears above? Crown on top?



My filter idea: _____



Which landmarks I'd use: _____



Save **screenshot** to your **portfolio folder**!





Mini-Project · Design a CV System!



Mini-Project · End-to-End CV Design



Mission: Design (on paper) a complete CV system for a real problem!

Step 1 · Pick a problem:



Spot tigers in
jungle photos



Help blind
read books



Detect helmet
on bikers

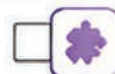
Step 2 · Define the AI task:



Classification



Detection



Segmentation

Step 3 · Data plan:



Need _____ images per class



Source of images: _____

Step 4 · Tool:



Teachable
Machine



Python+
TensorFlow



Roboflow

Step 5 · Where deployed:



Mobile app



Web app



Smart camera



Drone

Step 6 · Success metric:



Target accuracy: _____ %

Why? _____



Bias in Computer Vision



 **CV is biased when data is biased!** 



Real Bias Cases:



CASE 1: Soap Dispenser (2017)

An automatic dispenser at hotel didn't work for dark-skinned guests! Trained on light-skinned hands only!



CASE 2: Joy Buolamwini's Study (MIT)

She found face recognition: 99% accurate on white men, only 65% on dark-skinned women! Major fix needed!



CASE 3: Indian Saree Recognition

Western-trained AI saw sarees as "bedsheets" or "curtains"! Indian devs had to retrain with Bharat data!



How engineers FIX bias:

- Diverse training data (gender, skin, age, region)
- Test on multiple groups
- Include Indian context (clothing, food, places)
- Continuous monitoring



Fair AI starts with **FAIR** data!



▲ Ethics of CV · Surveillance!



CV gives power to watch – but who should have it?



The big questions:

-  Should cameras watch you in school?
-  In public streets? Always?
-  In offices?
-  In your home (Alexa cameras)?
-  In shops to track shopping?

Real cases:



CHINA

Has 700+ million CCTV cameras with face recognition!
Tracks every citizen's movement! Many call it concerning!



EUROPE

EU AI Act 2024 BANS most public face recognition!
Strong privacy laws!



INDIA

Delhi has 1.4 lakh CCTV cameras (highest density worldwide!). DigiYatra at airports uses face.
DPDP Act 2023 limits!

Your opinion (discuss with family):

CV in schools? ☐ Yes ☐ No ☐ With limits – Why? _____



Brain Teasers • CV!



CV Brain Buster!



Puzzle 1: A 200×200 colour photo has how many numbers?

Hint: each pixel = 3 numbers (RGB)

Answer: _____



Puzzle 2: Match CV task to use case:

A) Tumour boundary – _____

B) Tag friends in photo – _____

C) Count cars in image – _____

Options: Classification / Detection / Segmentation



Puzzle 3: Indian AI Qure.ai screens TB in 30 sec. Old way took 2 days. How many times faster?

2 days = 172,800 sec. $172,800 \div 30 = ?$

Answer: _____ times faster!



Puzzle 4: Pixel pattern: What colour is RGB (128, 0, 128)?

Answer: _____ (Hint: medium red + medium blue!)



Puzzle 5: A self-driving car processes camera 60 times/second. In 1 minute, how many image analyses?

Answer: _____

CV Vision Challenge!

1. **Pixel Hunt:** Take 1 photo. Count all the COLOURS – how many unique you see?
2. **Teachable Machine Build:** Do the portfolio project on page 17!
3. **Deepfake Detective:** Watch 3 viral videos. Spot any signs of AI?
4. **Snapchat Inspector:** Try 5 filters. Notice WHICH facial landmarks they target!
5. **CV in Daily Life:** Find 10 CV uses around you (cameras, phones, etc.)!
6. **Ethics Debate:** Discuss with family – should police use face recognition?



In Chapter 1 you learned:

- How AI sees PIXELS (RGB)
- **Classification, Detection, Segmentation**
- CV in medicine, cars, India!
- **Generative AI** (DALL-E, MidJourney)
- **Deepfakes** & how to spot them
- **Bias & ethics** in CV
- Built your **FIRST** AI model!



CV Specialist!



Next: NLP Mastery!

