

photomath



micro**blink**

Real-time OCR na mobilnim uređajima

Nenad Mikša, 13. 10. 2016.

Pojmovi

1. Real-time process

- proces koji se odvija u stvarnom vremenu

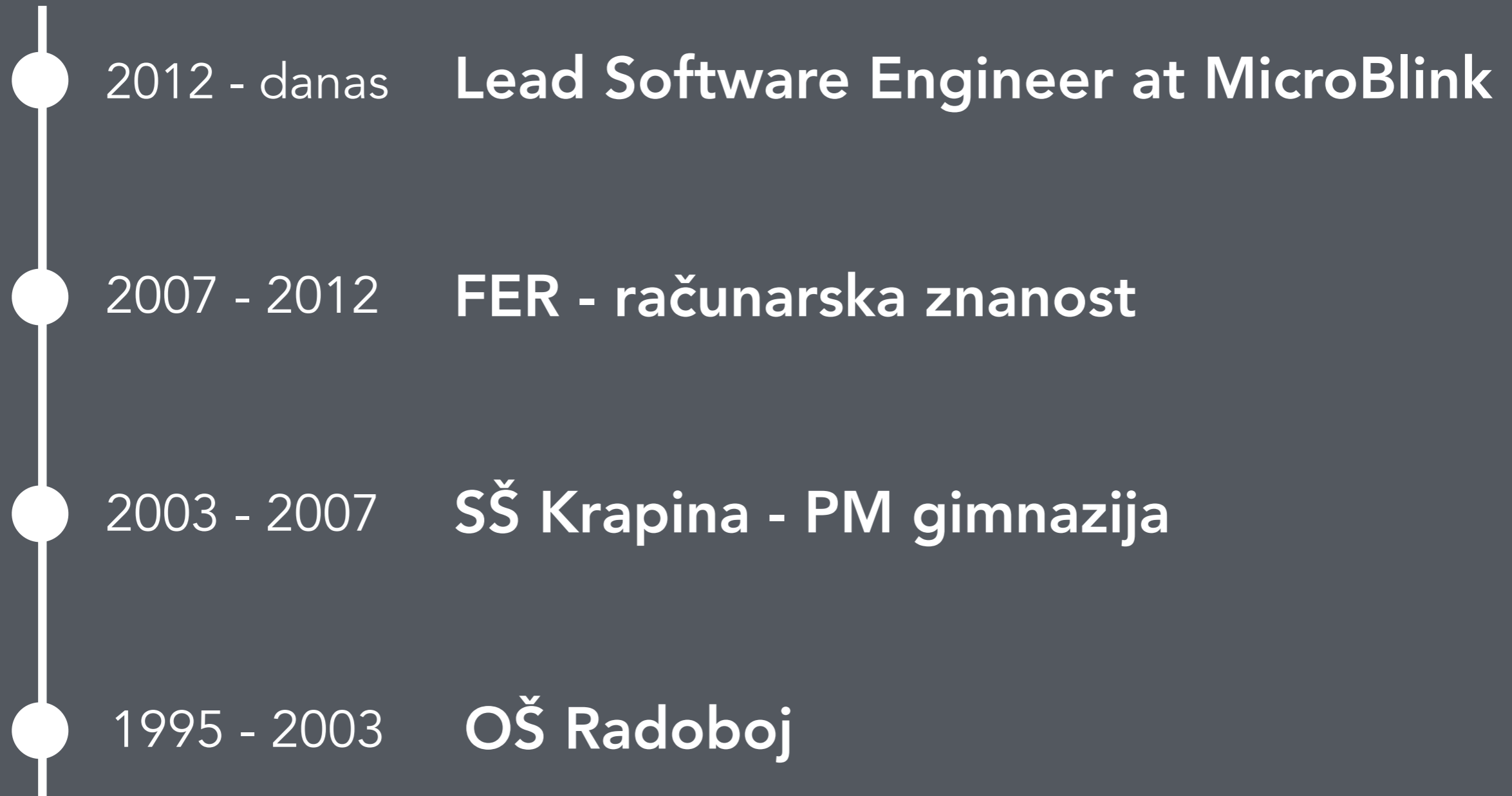
2. OCR (optical character recognition)

- optičko prepoznavanje znakova

3. Machine vision

- analiza slike s ciljem da se mogu donijeti neke odluke na temelju te slike

Nenad Mikša





microblink

mobile vision software



**Machine vision tehnologija
za mobilne uređaje**



photomath



Instant rezultat za skenirani matematički zadatak
















2 puta na vrhu App storea

iTunes Charts

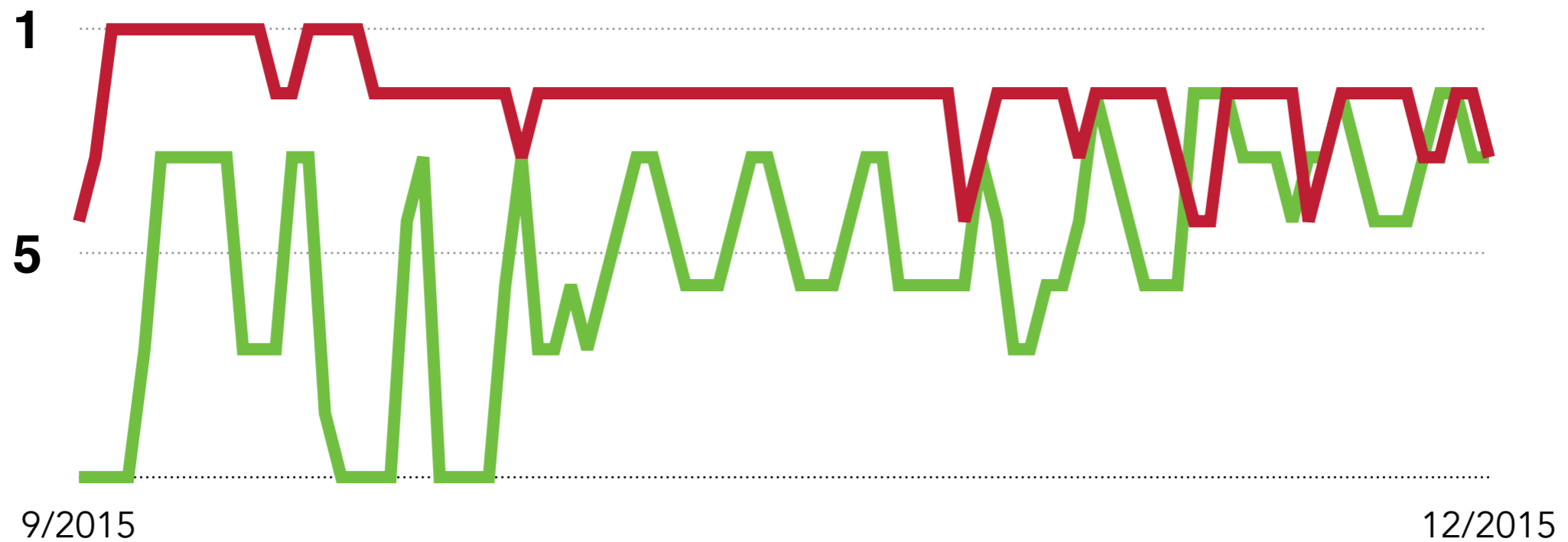
New content arrives on iTunes all the time. Here you can see what's new this week and browse the top 100 songs, albums, TV shows, movies, apps, and more.

[Shop the iTunes Store now >](#)

Featured Songs Albums **Free Apps** Paid Apps Books Movies TV Shows Music Videos

1.  PhotoMath Education Buy Now on iTunes >	2.  Messenger Social Networking Buy Now on iTunes >	3.  Happy Wheels Games Buy Now on iTunes >	4.  Instagram Photo & Video Buy Now on iTunes >	5.  Facebook Social Networking Buy Now on iTunes >
6.  YouTube Photo & Video Buy Now on iTunes >	7.  Snapchat Photo & Video Buy Now on iTunes >	8.  Layout from Photo & Video Buy Now on iTunes >	9.  Pandora Radio Music Buy Now on iTunes >	10.  Google Maps Navigation Buy Now on iTunes >
11. 	12. 	13. 	14. 	15. 

App Store rank (Education)



- Photomath
- Duolingo (za usporedbu)

Active Users

Apr 7, 2016 - May 4, 2016

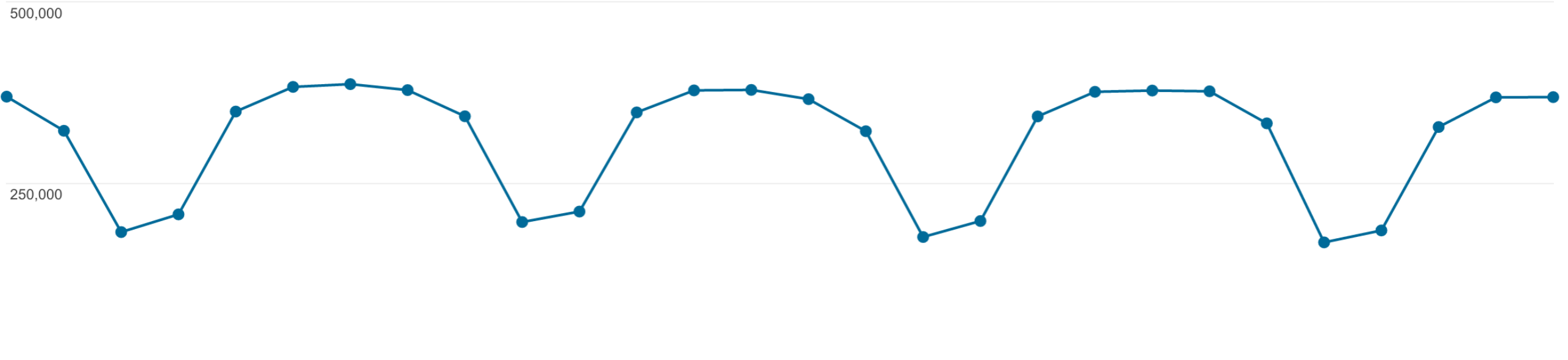
Email Export Add to Dashboard Shortcut

All Users
100.00% Users

+ Add Segment

Active Users

1 Day Active Users 7 Day Active Users 14 Day Active Users 30 Day Active Users



Apr 8

Apr 15

Apr 22

Apr 29

1 Day Active Users

368,454

% of Total: 100.00% (368,454)

7 Day Active Users

1,470,968

% of Total: 100.00% (1,470,968)

14 Day Active Users

2,393,662

% of Total: 100.00% (2,393,662)

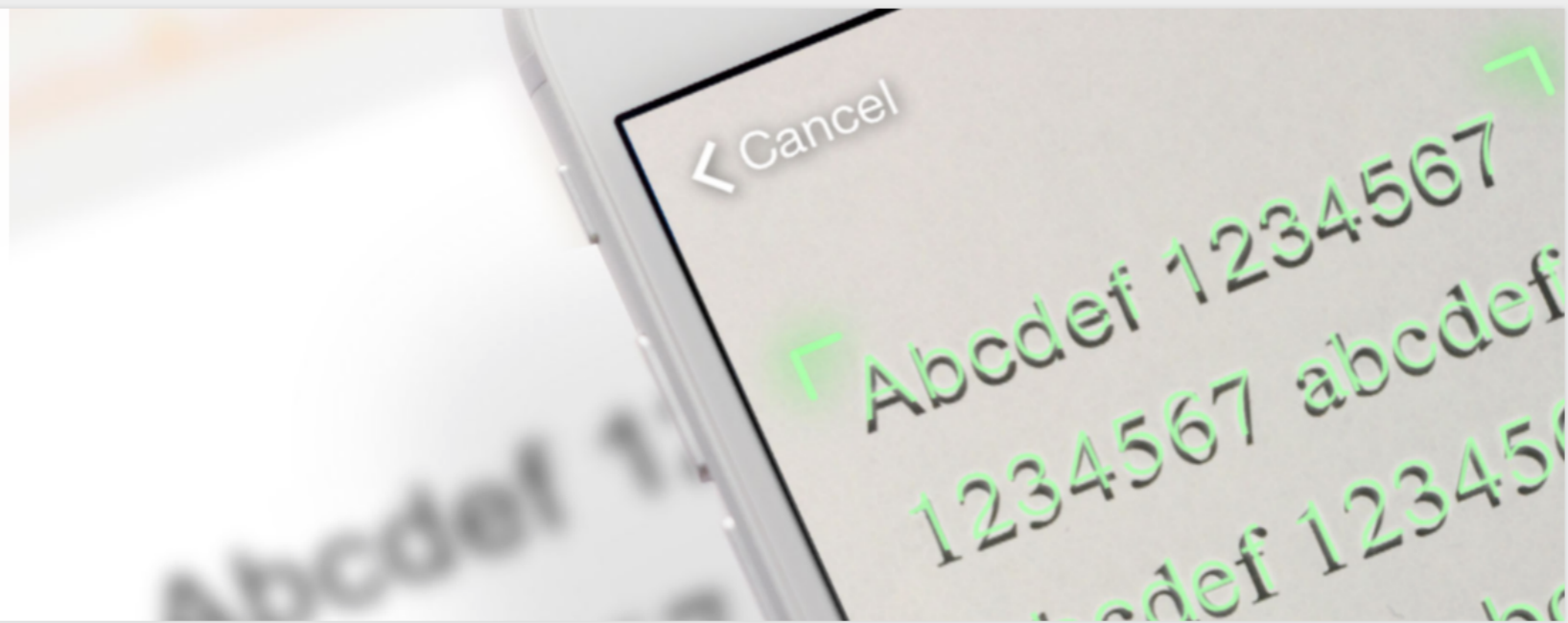
30 Day Active Users

4,109,062

% of Total: 100.00% (4,109,062)

blinkOCR

Real-time camera text recognition SDK for mobile apps



blinkID

Real-time ID card, driver license & passport scanning SDK for mobile devices



photopay

Mobile payment data capture SDK for mobile banking and payment apps



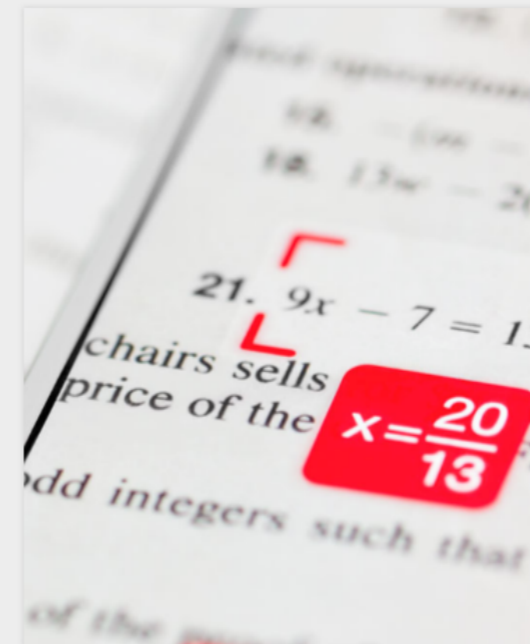
blinkbarcode

High performance 1D & 2D barcode scanning SDK



photomath

The world's smartest camera calculator

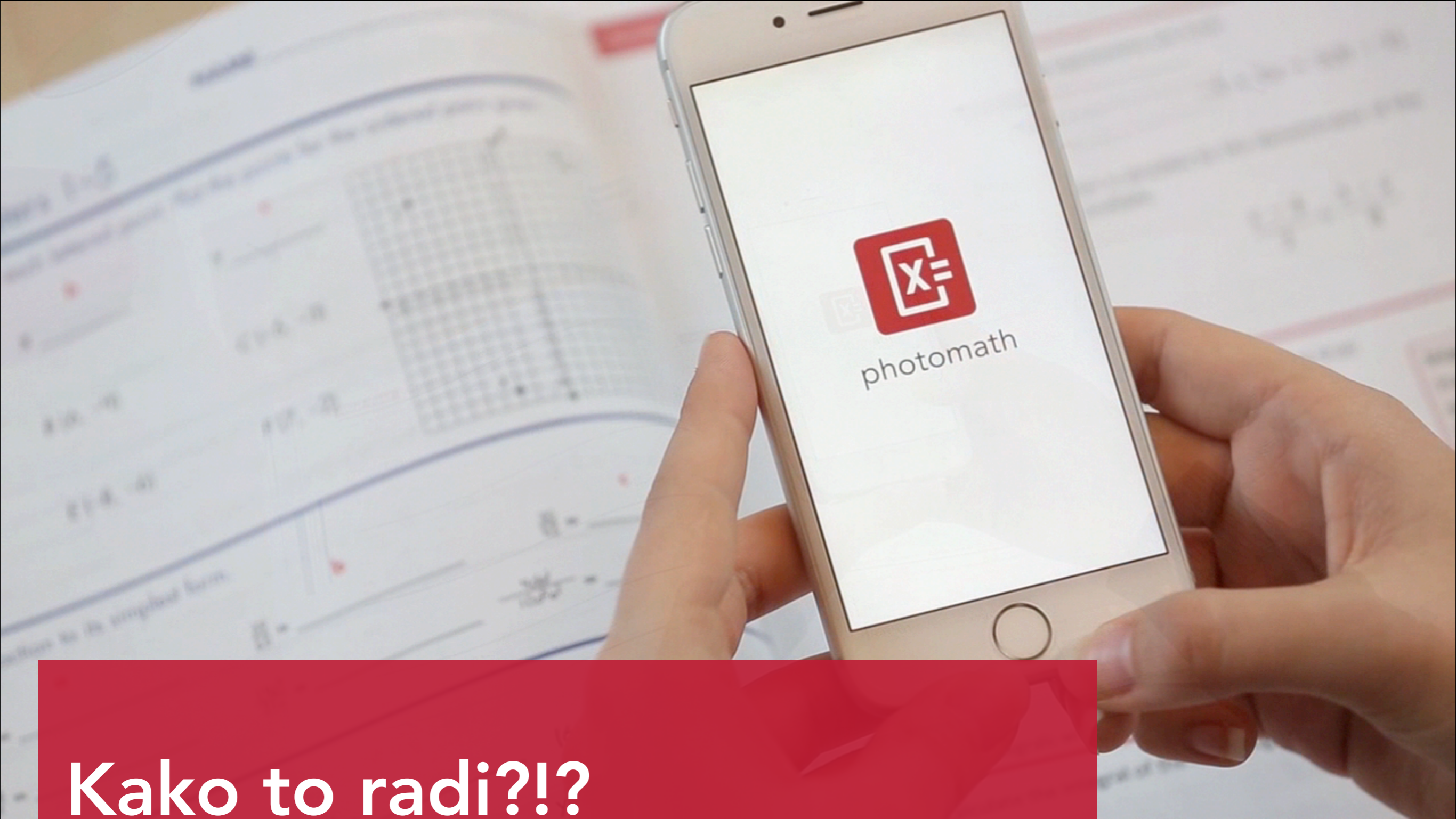




General Electric







Kako to radi?!?

Proces obrade

$$\begin{array}{rcllcl} 4x & - & 3y & + & z & = & 2 \\ -2x & - & y & - & 2z & = & 2 \\ -x & + & 2y & + & 4z & = & -9 \end{array}$$

Rezolucija

945x272

257.040 pixela

1. **Predobrada** slike (image processing)
2. **Detekcija** znakova (object detection)
3. **Klasifikacija** znakova (machine learning)
4. **Ekstrakcija** matematičkog izraza
5. **Rješavanje** matematičkih izraza (ekspertni sustav)

Ulazna slika

$$\begin{array}{rcccccc} 4x & - & 3y & + & z & = & 2 \\ -2x & - & y & - & 2z & = & 2 \\ -x & + & 2y & + & 4z & = & -9 \end{array}$$

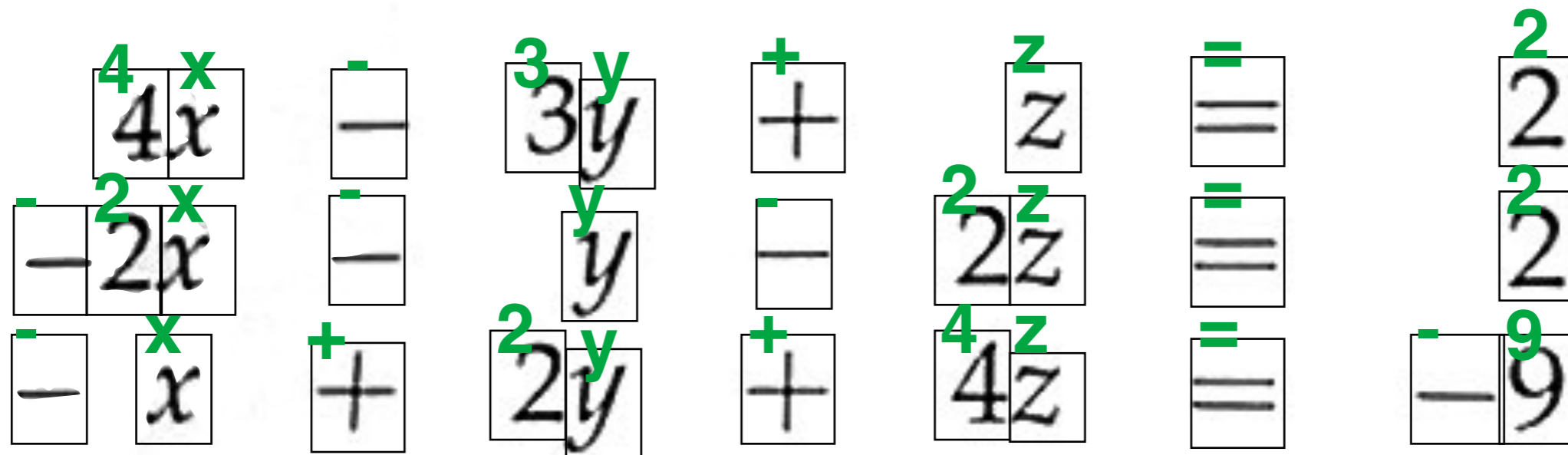
1. Predobrada slike

$$\begin{array}{rcccccc} 4x & - & 3y & + & z & = & 2 \\ -2x & - & y & - & 2z & = & 2 \\ -x & + & 2y & + & 4z & = & -9 \end{array}$$

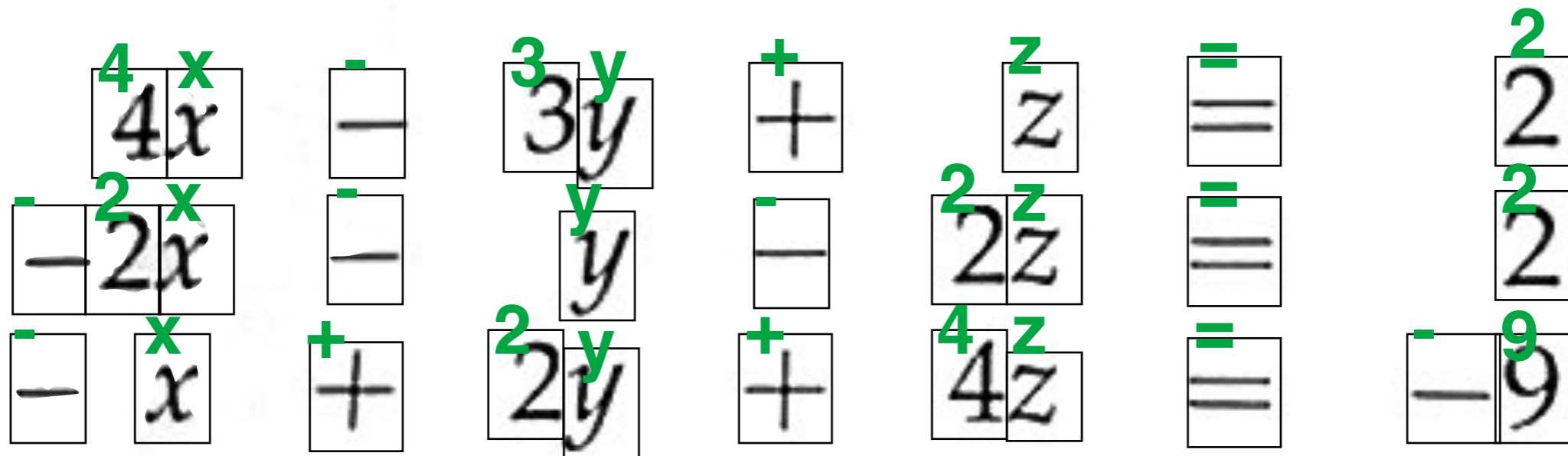
2. Detekcija znakova

$$\begin{array}{r} 4x \\ -2x \\ -x \end{array} \begin{array}{c} - \\ - \\ + \end{array} \begin{array}{r} 3y \\ y \\ 2y \end{array} \begin{array}{c} + \\ - \\ + \end{array} \begin{array}{r} z \\ 2z \\ 4z \end{array} \begin{array}{c} = \\ = \\ = \end{array} \begin{array}{r} 2 \\ 2 \\ -9 \end{array}$$

3. Klasifikacija znakova



4. Ekstrakcija matematičkog izraza



system(

 equals(add(sub(mul(4,x),mul(3,y)),z),2),

 equals(sub(sub(mul(neg(2),x),y),mul(2,z)),2),

 equals(add(add(neg(1),x),mul(2,y)),mul(4,z)),neg(9))

)

5. Rješavanje izraza

◀ Back to App Store 10:29 97%
 < Back Solving Steps

$$\begin{cases} 4x - 3y + z = 2 \\ -2x - y - 2z = 2 \\ -x + 2y + 4z = -9 \end{cases}$$

Sum it with $-2x - y - 2z = 2$ multiplied by 2

$$\begin{cases} 4x - 3y + z = 2 \\ -2x - y - 2z = 2 \\ (-1 + 2 \times -2)x + (2 + 2 \times -1)y + (4 + 2 \times -2)z = -9 + 2 \times 2 \end{cases} \quad (1)$$

$$\begin{cases} 4x - 3y + z = 2 \\ -2x - y - 2z = 2 \\ (-1 + 2 \times -2)x + (2 - 2)y + (4 + 2 \times -2)z = -9 + 2 \times 2 \end{cases} \quad (2)$$

$$\begin{cases} 4x - 3y + z = 2 \\ -2x - y - 2z = 2 \\ (-1 + 2 \times -2)x + 0y + (4 + 2 \times -2)z = -9 + 2 \times 2 \end{cases} \quad (3)$$

Next step →

◀ Back to App Store 10:30 97%
 < Back Solving Steps

$$x = 1$$

Substitute y with 0 because $y = 0$

$$\begin{cases} -3 \times 0 + z = -2 \\ y = 0 \\ x = 1 \end{cases} \quad (37)$$

$$\begin{cases} 0 + z = -2 \\ y = 0 \\ x = 1 \end{cases} \quad (38)$$

$$\begin{cases} z = -2 \\ y = 0 \\ x = 1 \end{cases}$$

← Previous step Next step →

was Photomath helpful to you?

$$25x^2 - \quad = 0$$

~~$$b) \sqrt[3]{27.5}$$~~

~~$$\therefore \sqrt[3]{3.5} = 3i$$~~

~~$$\sqrt[3]{27x^6} =$$~~

$$x^3 - 6x^2 + 13x - 8 = 0$$

~~$$2x + 3 = 0$$~~

$$\lim_{x \rightarrow 0} \frac{(\arctan x)^2}{\ln(1+x^2)}$$

$$2x + 5 = 11$$

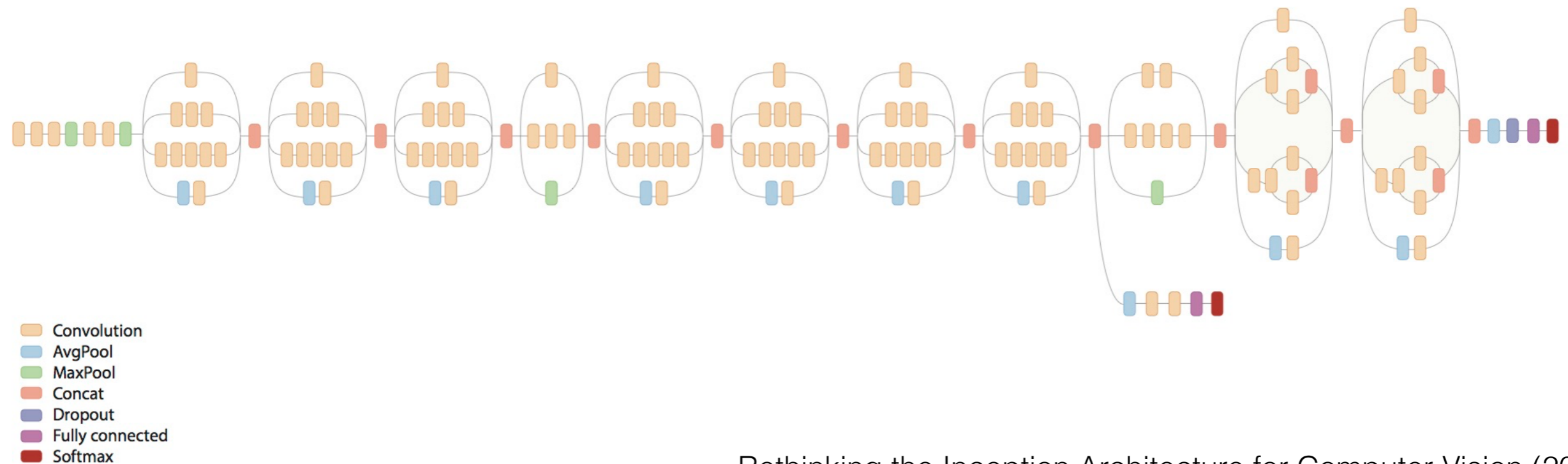


← Klasični pristup

Budućnost -
deep learning →

Inception model

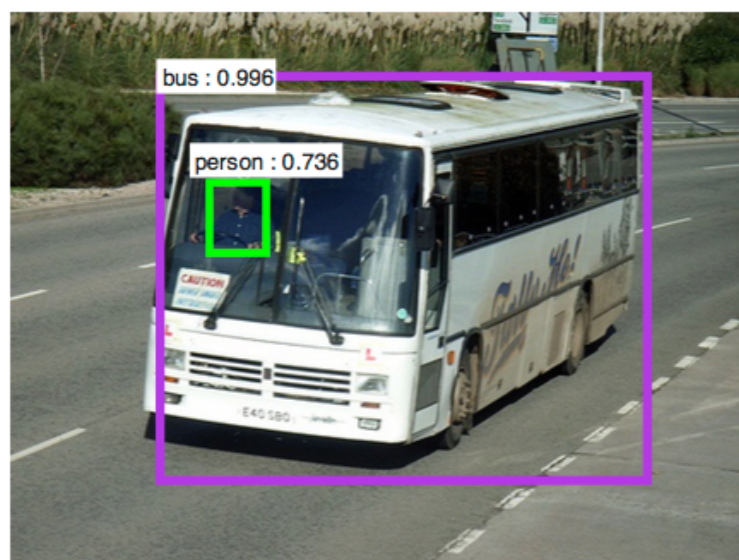
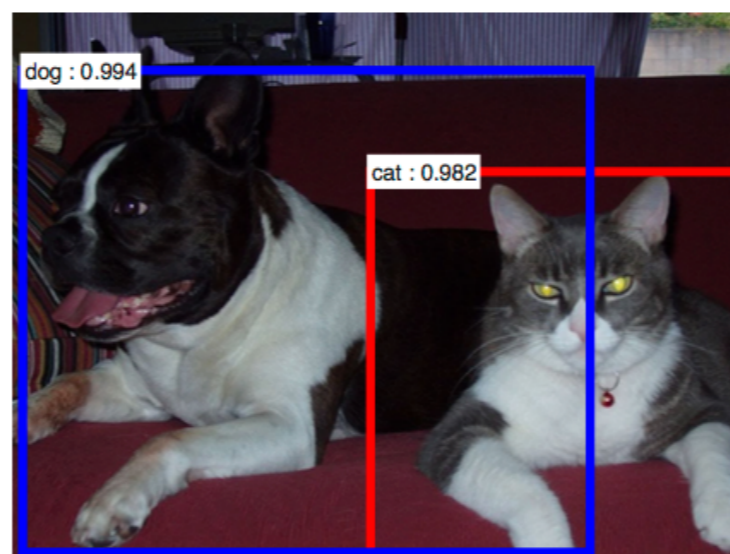
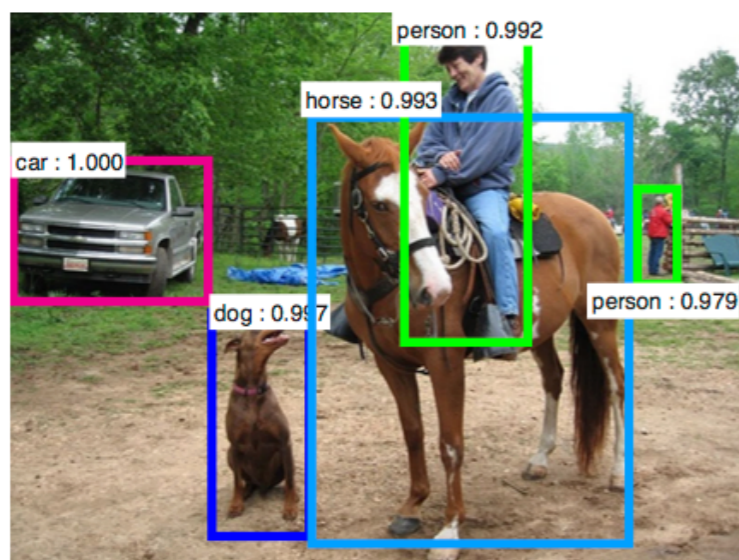
State of the art model za
generalni image recognition



Rethinking the Inception Architecture for Computer Vision (2015)
Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna

Faster R-CNN model

Istodobno detekcija i klasifikacija



Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks (2015)
Ren, Shaoqing and He, Kaiming and Girshick, Ross and Sun, Jian

Photomath Dataset

- **5 milijuna** slika jednadžbi = **50 milijuna** slika znakova
- Potencijal za skupljanje **500k slika dnevno**
- Ručno labeliramo **100 tisuća slika znakova tjedno**
- **Potencijalno najveći dataset** za OCR na svijetu.

$$\frac{-5x+9}{2x^2} dx$$

$$\sqrt{11 \cdot 2^2} \cdot \sqrt{11 \cdot 3^4}$$

$$\frac{1}{x^2+x+1} - \frac{x^2+2x}{(x^3-1)} - \frac{2x+2}{1-x^2}$$

$$\frac{{}^4\log 9 \cdot {}^3\log 2 + {}^4\log 8}{{}^9\log 6 - {}^9\log 2} = \dots$$

$$\sqrt{11 \cdot 2^2} \cdot \sqrt{11 \cdot 3^4}$$

$$+\frac{2}{1-x}$$

$$13+2x-98 = 5x+83-7x$$

611
+ 155

$$36a + 5c - 4ac - 45$$

$$\frac{2x+4}{x^3+x^2-2x} + \frac{2}{1-x}$$

a sečtete:

$$+ 3c)(-4abc)$$

$$10 + \frac{8}{4} - 4$$

$$\frac{0.04 \times 14 \times 0.05}{30 \times 0.035 \times 0.8}$$

$$\lg^2 5 + (\lg 5 + 2) \lg 20$$

$$a / x + 4 = 5$$

ky

$$x^2 - 10x + 25 < 0$$

$$4+7=$$

Hasil dari $3\frac{3}{4} : 2\frac{1}{2} - 4\frac{1}{3}$ adalah

$$3x + 7 = 43$$

$$-3)(2x+3) - 4(x+2)$$

$$2x + y = 12$$

$$\text{Hasil dari } \frac{{}^4\log 9 \cdot {}^3\log 2 + {}^4\log 8}{{}^9\log 6 - {}^9\log 2} = \dots$$

$$-1,5x + 9$$

$$\frac{{}^4\log 9 \cdot {}^3\log 2 + {}^4\log 8}{{}^9\log 6 - {}^9\log 2}$$

$$2 + 2 * 3$$

$$10(2) - 4$$

$$2x - 1 = 3$$

$$-\frac{3}{4} \cdot \left(\frac{4}{5} + \frac{2}{3}\right) =$$

$$7\frac{3}{4} \div 2 =$$

$$\lim_{x \rightarrow \infty} (1+x^6)^{\frac{1}{x^4 \ln^2 3x}}$$

$$3 + \left[\frac{x-3y}{2} - \left(\frac{1-x}{2} + \frac{y+6}{4}\right)\right] = 2x+1$$

$$\frac{x+1}{x-1} = \frac{x-1}{x+1}$$

$$7 - 2 + 6 - 2 - 8 =$$

$$2 \cdot 3x + 5x$$

$$3x + 2x =$$

$$5 \cdot 6 + 3 : 2$$

$$2 + 2 =$$

$$\left(\frac{1}{3}\right)(2x+3) - (2x-\frac{1}{3}) = \frac{3x-2}{6}$$

$$10 - 5$$

$$10 + 5$$

$$1000000 + -800000000000$$

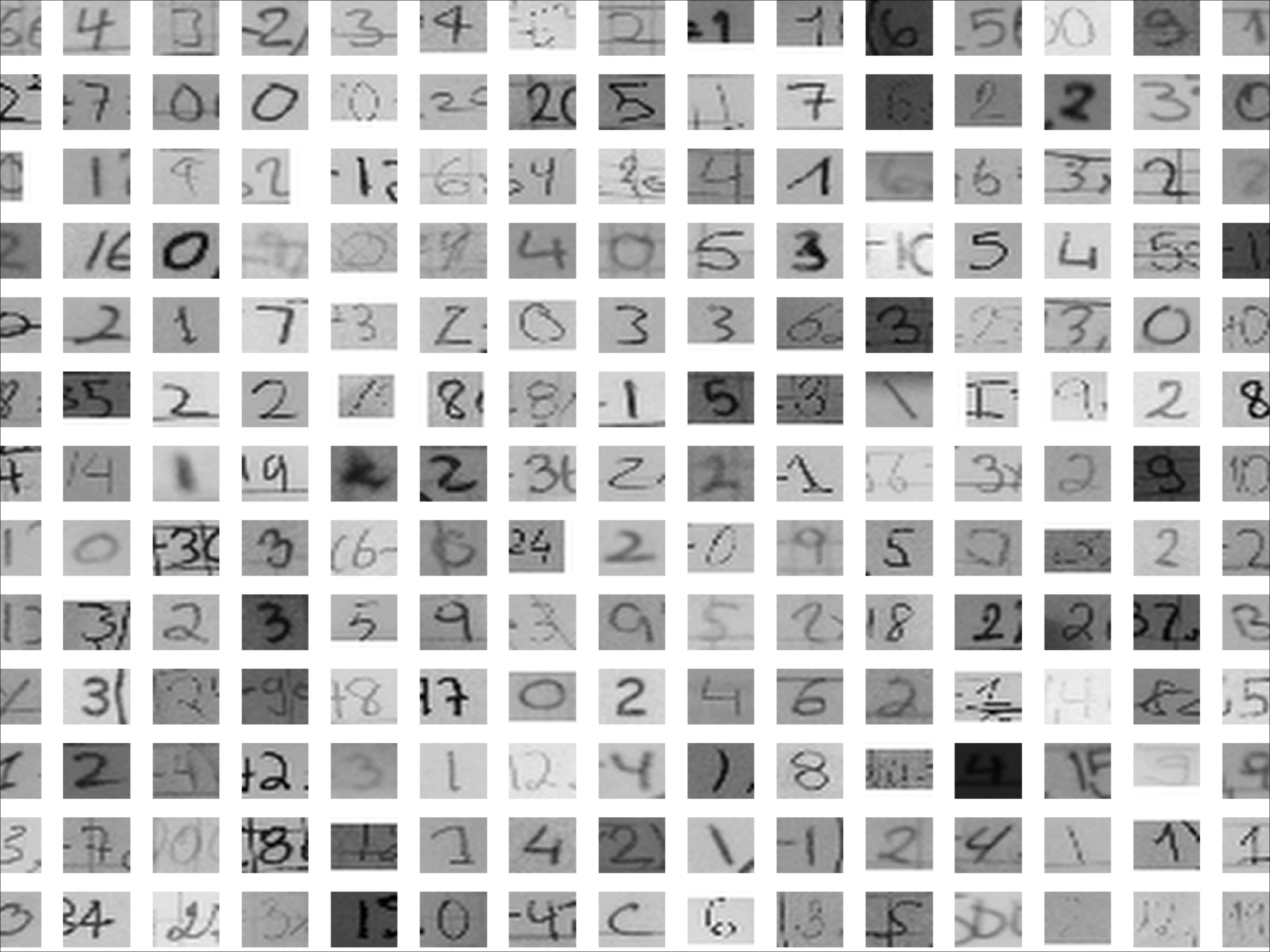
$$(1+x)^{1/x} = ?$$

$$(6x^3 - 15x^2 + 10x - 2) : (3x^2 - 6x + 2) = 2x + 1$$

$$9 = 3 + 6$$

$$\frac{2-x-1}{3} = 0$$

$$2x(x$$





GEFORCE GTX

GIGABYTE

GEFORCE GTX

GIGABYTE

GEFORCE GTX


GIGABYTE

GEFORCE GTX

GIGABYTE

Cilj:

- nova generacija OCR enginea
- za Photomath i sve druge proizvode MicroBlinka

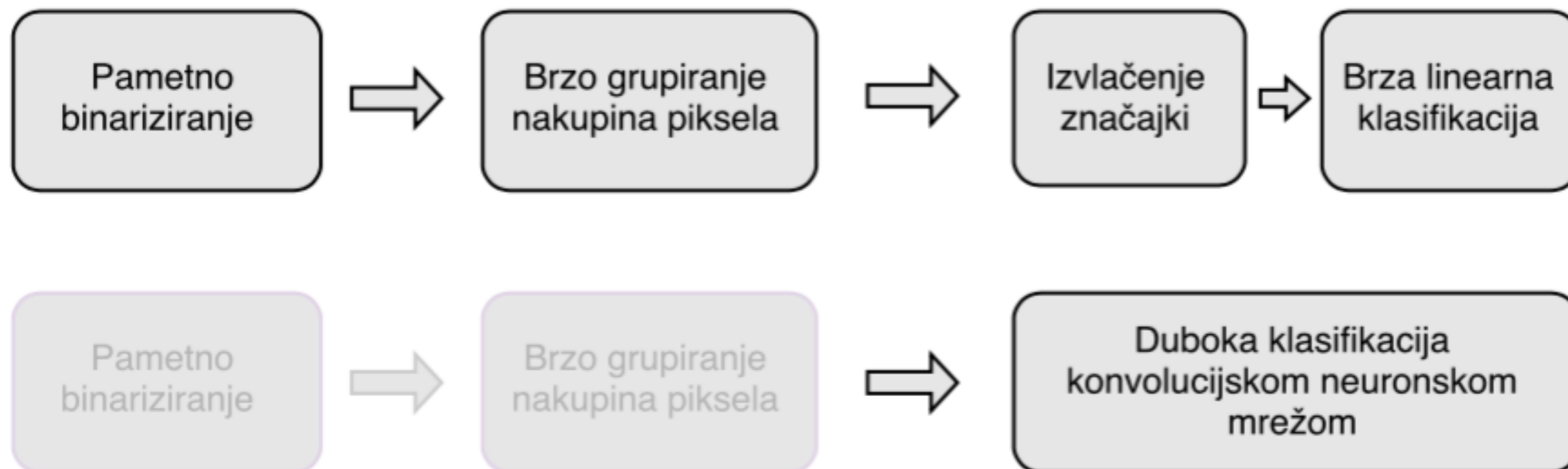
Light 

Abcdef 1234567
1234567 abcdef
Abcdef 12345
1234567 ab

Predobrada slike

Detekcija znakova

Klasifikacija znakova



Duboka detekcija i klasifikacija sa specijaliziranom OCR neuronskom arhitekturom

Ulaz originalna slika --> Izlaz lokacije i klase svih znakova



microblink

mobile vision software



**Sve to treba raditi u
stvarnom vremenu**

Brzina je ključna

- svi opisani koraci se moraju provesti u manje od 1 sekunde na slabim uređajima
- odnosno manje od 0,1 s na jakim uređajima

Brzina je ključna (2)

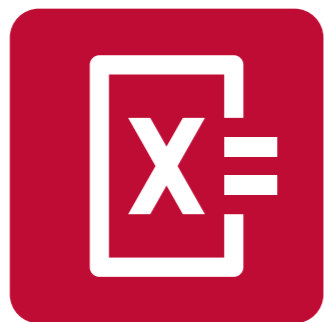
- algoritmi niske složenosti
- C++14
- SIMD assembly
 - single instruction - multiple data
- GPU
 - OpenCL, OpenGL, Vulkan, Metal, DirectX

Ali i jednostavnost!

- jednostavan API za korisnike
- pristupačna i čitljiva dokumentacija
- primjeri integracijskog koda
- Java wrapper za Android
- Objective C / Swift wrapperi za iOS
- SDK javno dostupan na GitHub-u

Savjeti

- natječite se iz informatike i matematike
- učite algoritme
- učite C++, pogotovo inačice C++14 i novije
 - svladajte CMake build system
- svladajte Git
- sudjelujte u open source projektima
- svladajte i (barem) jedan VM jezik, npr. Javu
- skriptirajte (Bash, Python)



photomath



micro**blink**

Photomath, Inc.
MicroBlink, Ltd.

nenad.miksa@microblink.com

Hvala na pažnji!