



# Gestų kalbos atpažinimas naudojant internetinę kamerą

Pranciškus Ambazas

VU MIF

Programų sistemų katedra

*2017-06-26*



# Darbo tikslas

Išsiaiškinti, ištirti ir išanalizuoti galimybes gestų kalbą internetinės kameros pagalba versti į rašytinę ar žodinę kalbą.



# Uždaviniai

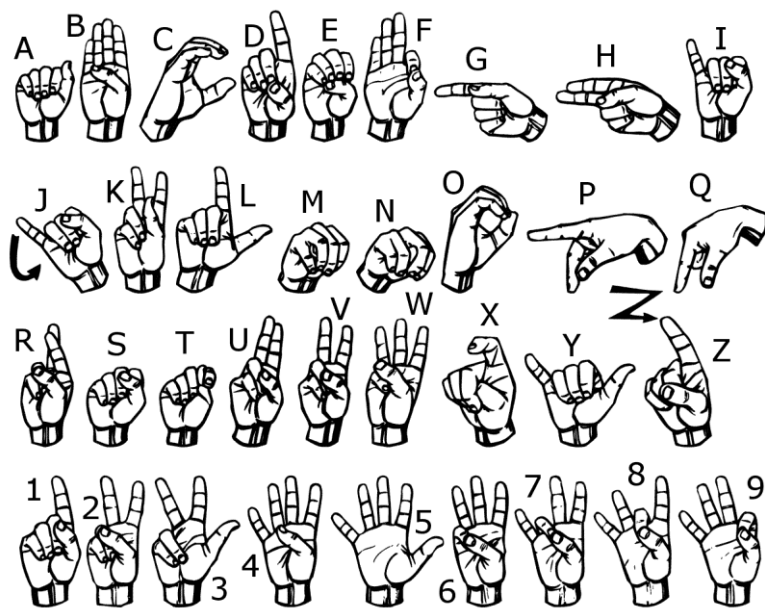
- Išsiaiškinti gestų kalbos ypatybes;
- Iširti galimybes atpažinti gestų kalbą;
- Išbandyti skirtingus būdus atpažinti statinę gestų kalbą.



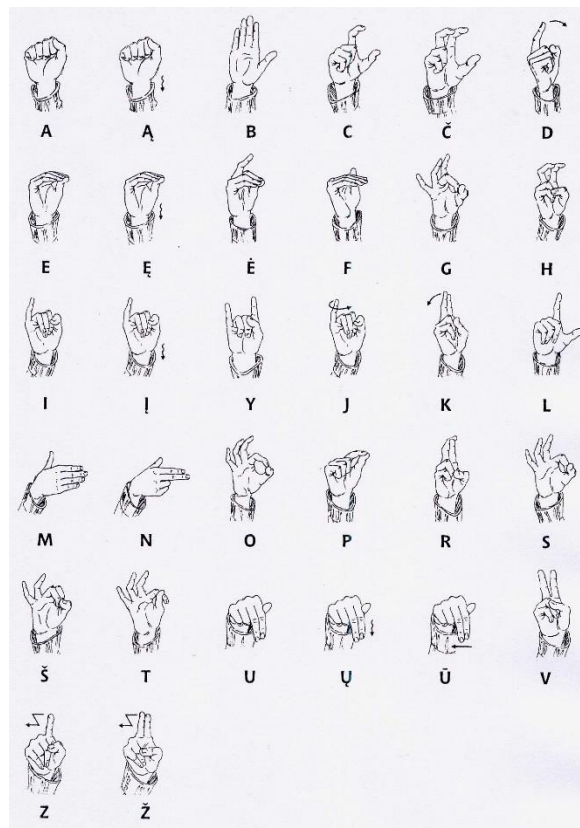
# Gestų kalba

- 360 milijonų žmonių su klausos problemomis;
- 32 milijonai iš jų – vaikai;
- ASL – populiariausia kalba po „didžiojo ketverto“;
- Statiniai ir dinaminiai judesiai.

# Gestų kalba



Amerikiečių anglų gestų kalbos abėcėlė



Lietuvių gestų kalbos abėcėlė



# Bandymų duomenys

- 24-ių ASL abėcėlės raidžių gestai;
- Kiekvienai raidei – katalogas;



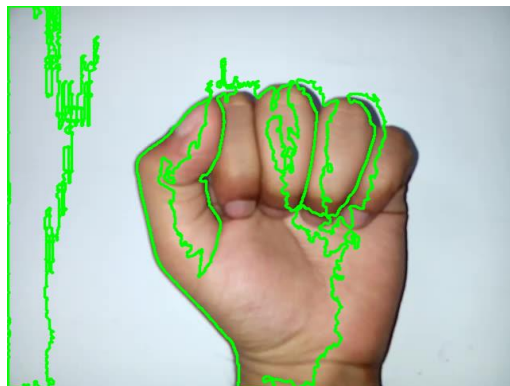
# Įprastiniai metodai

- OpenCV;
- Trukmė:  $\sim 5$  valandas;
- 400 kadro kiekvienai ASL raidei;
- 95-98% tikslumas po apmokymo;

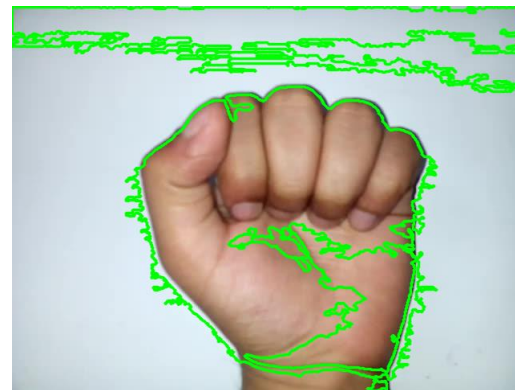
# Kadro konvertacija



Originalus paveikslėlis



Pritaikyta  $G_x$



Pritaikyta  $G_y$



Pritaikyta  $G$



Be fono

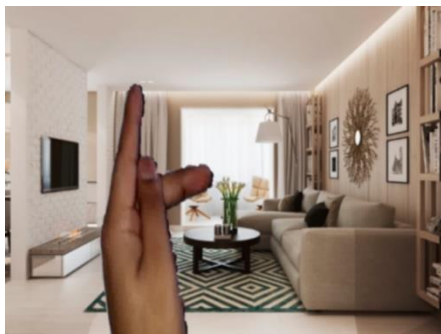


Dviejų spalvų



# Konvoliucinio tinklo modelis

- 1+5 bandymai;
- Skirtingi apmokymai;
- Skirtingi duomenų kiekiai;



K raidės gestas



K raidės gestas



# Konvoliucinio tinklo modelis.

## Nuodugnus apmokymas

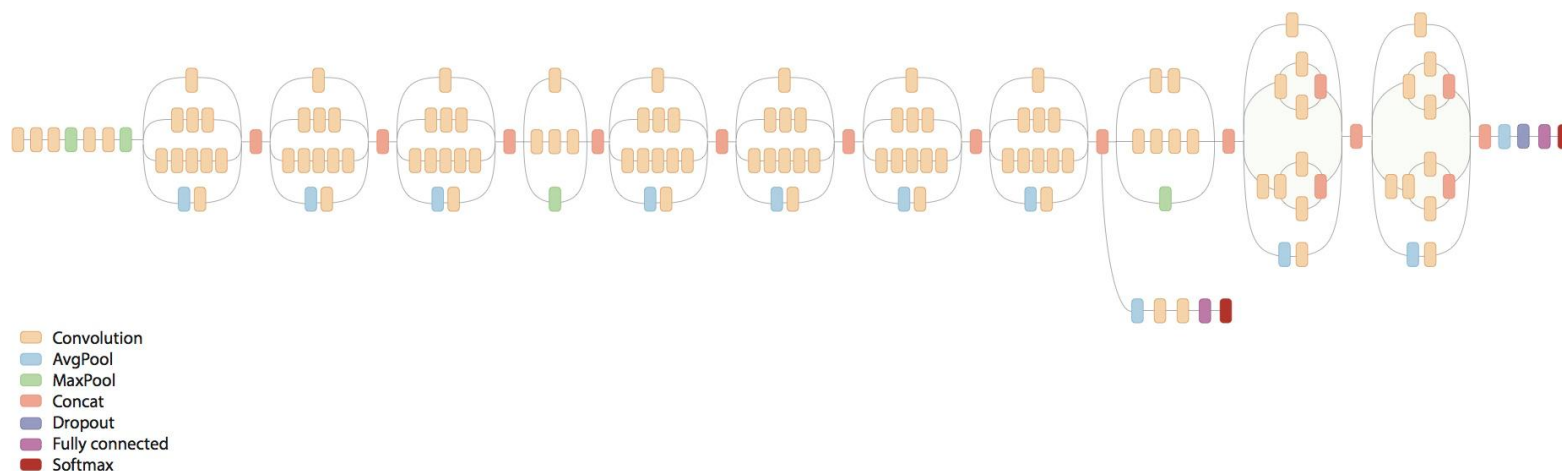
- Pasinaudota VGG16 architektūra;
- Dvi raidės: A ir B;
- **Trukmė:** 2,5 h.;
- **Duomenys:** 400 kadro mokymui ir 200 pasitikrinimui kiekvienai raidei;
- **Rezultatas:** 50-60%.



# Konvoliucinio tinklo modelis.

## Dalinis apmokymas

- „Inception-v3“ modelis





# Konvoliucinio tinklo modelis.

## Pirmas bandymas

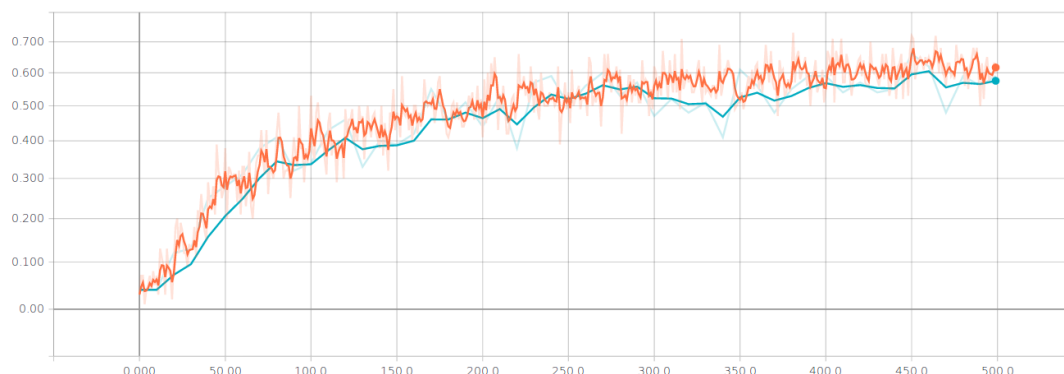
- Dvi raidės – A ir B;
- **Duomenys:** 400 kadro kiekvienai raidei;
- **Trukmė:**  $\sim 5$  min;
- **Rezultatas:** 95-98%.



# Konvoliucinio tinklo modelis.

## Antras bandymas

- **Duomenys:** 400 kadru kiekvienai raidei;
- **Trukmė:**  $\sim 30$  min;
- **Žingsniai:** 500;
- **Rezultatas:**  $\sim 60\%$  ir  $\sim 60\%$ .

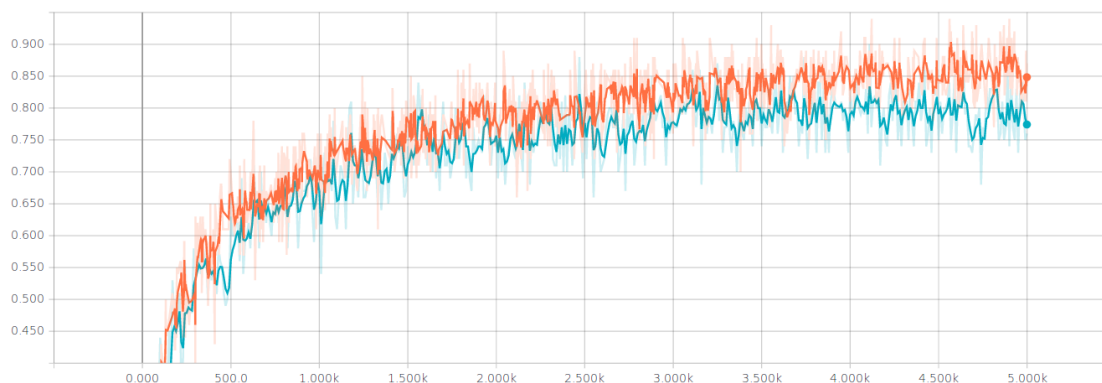




# Konvoliucinio tinklo modelis.

## Trečias bandymas

- **Duomenys:** 400 kadro kiekvienai raidei;
- **Trukmė:**  $\sim 30$  min;
- **Žingsniai:** 5000;
- **Rezultatas:** 85% ir 77%.

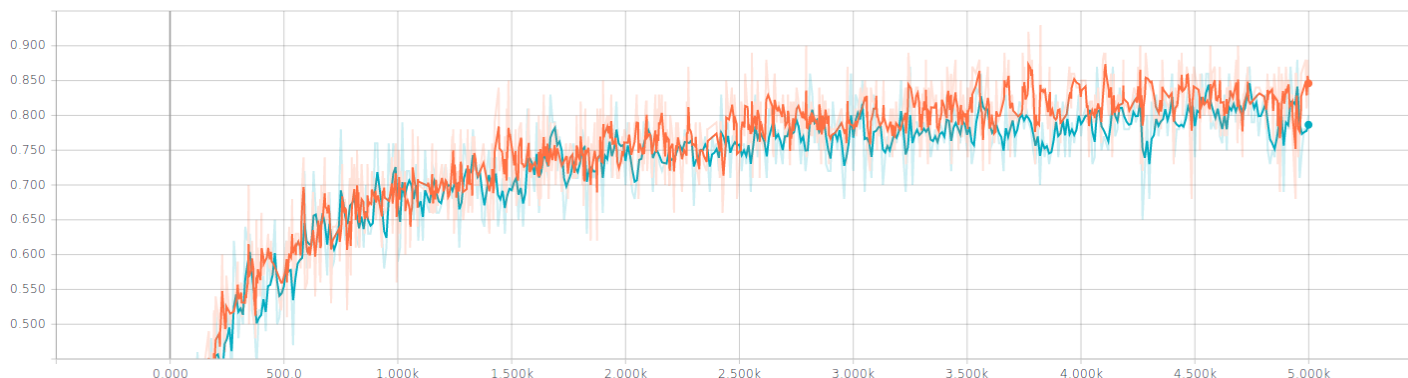




# Konvoliucinio tinklo modelis.

## Ketvirtas bandymas

- **Duomenys:** 1000 kadru kiekvienai raidei;
- **Trukmė:**  $\sim 45$  min;
- **Žingsniai:** 5000;
- **Rezultatas:** 85% ir 79%.

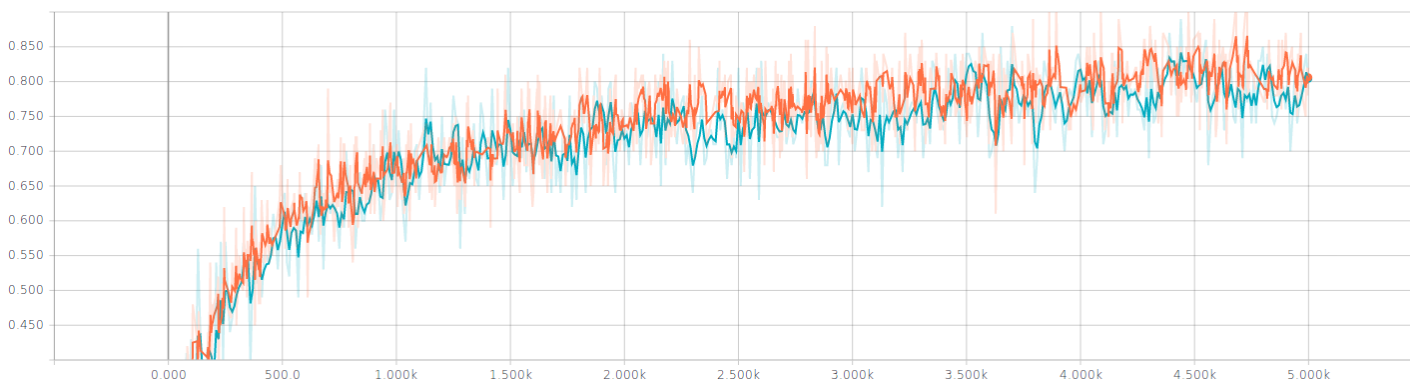




# Konvoliucinio tinklo modelis.

## Penktas bandymas

- **Duomenys:** 3000 kadru kiekvienai raidei;
- **Trukmė:**  $\sim 45$  min;
- **Žingsniai:** 5000;
- **Rezultatas:** 80% ir 80%.







# Rezultatai

1 lentelė. Konvoliucinių neuroninių tinklų ASL abėcėlės apmokymų statistika

	<b>Antras bandymas</b>	<b>Trečias bandymas</b>	<b>Ketvirtas bandymas</b>	<b>Penktas bandymas</b>
Apmokymo tikslumas	60%	85%	85%	80%
Patikrinimo tikslumas	60%	77%	79%	80%

2 lentelė. Testavimo statistika

	<b>Ketvirtas bandymas</b>	<b>Penktas bandymas</b>
Pirmas pasirinkimas	15,13%	14,71%
Vienas iš dviejų pasirinkimas	25,21%	20,17%
Vienas iš penkių pasirinkimas	38,24%	34,45%



# Rezultatai ir išvados

- Konvoliuciniai neuroniniai tinklai – galingas ir lengvai apmokomas įrankis;
- Dalinis apmokymas greičiau ir tiksliau už nuodugų;
- Didesni duomenų rinkiniai ir skirtingos duomenų bazės leistų padidinti tikslumą;
- Pilnai išvystytas modelis būtų ypač naudingas gestakalbiams ir bendrajančioms su jais.