

Detekcija analognog sata i prepoznavanje vremena

Upotrebom konvolutivnih neurnoskih mreža (CNN)

Filip Ivković SW33/2016, Katarina Tukelić SW58/2016



Uvod

U različitim industrijskim granama postoji česta potreba da se vrednosti određenih veličina očitavaju sa analognih interfejsa (*satovi*, *termometri*, *manometri itd.*) ručno, kako bi se kasnije te vrednosti obrađivale i analizirale, najčešće digitalno. To najčešće znači da ljudski operater mora da periodično obilazi, očitava, i unosi trenutne vrednosti u sistem za prikupljanje i obradu podataka kako bi ti podaci bili obrađeni. Rešavanjem ovog problema omogućila bi se integracija očitavanja vrednosti sa analognih interfejsa u automatizovane sisteme za prikupljanje i obradu podataka.

Cilj istraživanja

Cilj istraživanja je detekcija analognog sata na video snimku u realnom vremenu i očitavanje trenutnog vremena korišćenjem računarske vizije i mašinskog učenja. Konkretno, naš pristup za cilj ima to da se upotrebom *Konvolutivne Neuronske Mreže (CNN)* kreira model i istrenira se do te mere da može da izvrši detekciju analogng sata na video snimku u realnom vremenu, i da zatim prepozna trenutno vreme koje je na tom satu prikazano.

Algoritam

Za dostizanje našeg cija korišćene su dve CNN mreže. Zadatak prve neuronske mreže je da na video snimku, u realnom vremenu, detektuje analogni sat i, sa dovoljno velikom preciznošću, pronađe njegov *bounding box*. Sadržaj tog *bounding box-a* predstavlja ulaz u drugu mrežu, koja određuje koliko je sati i minuta prikazano na satu.

CNN za lokalizaciju sata

- Input za prvu mrežu je *grayscale* frame dimenzija *100x100*
- Neuronska mreža je standardna CNN sa 4 konvoluciona sloja i pooling slojem između svakog od njih
- Izlaz iz mreže je *multi-hot array* od 200 elemenata u kojem je naznačeno koji pikseli pripadaju okviru oko detektovanog sata

CNN za detekciju vremena na satu

- Sat se zatim iseca sa originalne slike, vrši se *resize* na dimenzije 100x100, i takva slika se prosleđuje drugoj neuronskoj mreži
- Osim po parametrima, ova neuronska mreža se razlikuje od prve po tome što se nakon konvolucija ona grana na dve grane, kako bi rešila problem detekcije sata i minuta odvojeno, na različite načine
- Prepoznavanje trenutnog broja sati se posmatra kao *klasifikacioni problem*, dok se detekcija broja minuta posmatra kao *regresioni problem*

Skup podataka

Za svaku mrežu generisan je poseban skup podataka na kojem je mreža trenirana i testirana. Oba seta su sačinjena od 12 tipova satova različitih veličina, rotiranih za određeni broj stepeni, koji prikazuju nasumično sate i minute. U skupovima podataka se takođe nalaze i slike na kojima nema sata.



Neki od korišćenih tipova sata

Slike korišćene za generisanje pozadine

Razlika između skupova je to što slike generisane za treniranje mreže za prepoznavanje vremena prikazuju satove mnogo većih razmera, tako da sat skoro u potpunosti prekriva sliku.



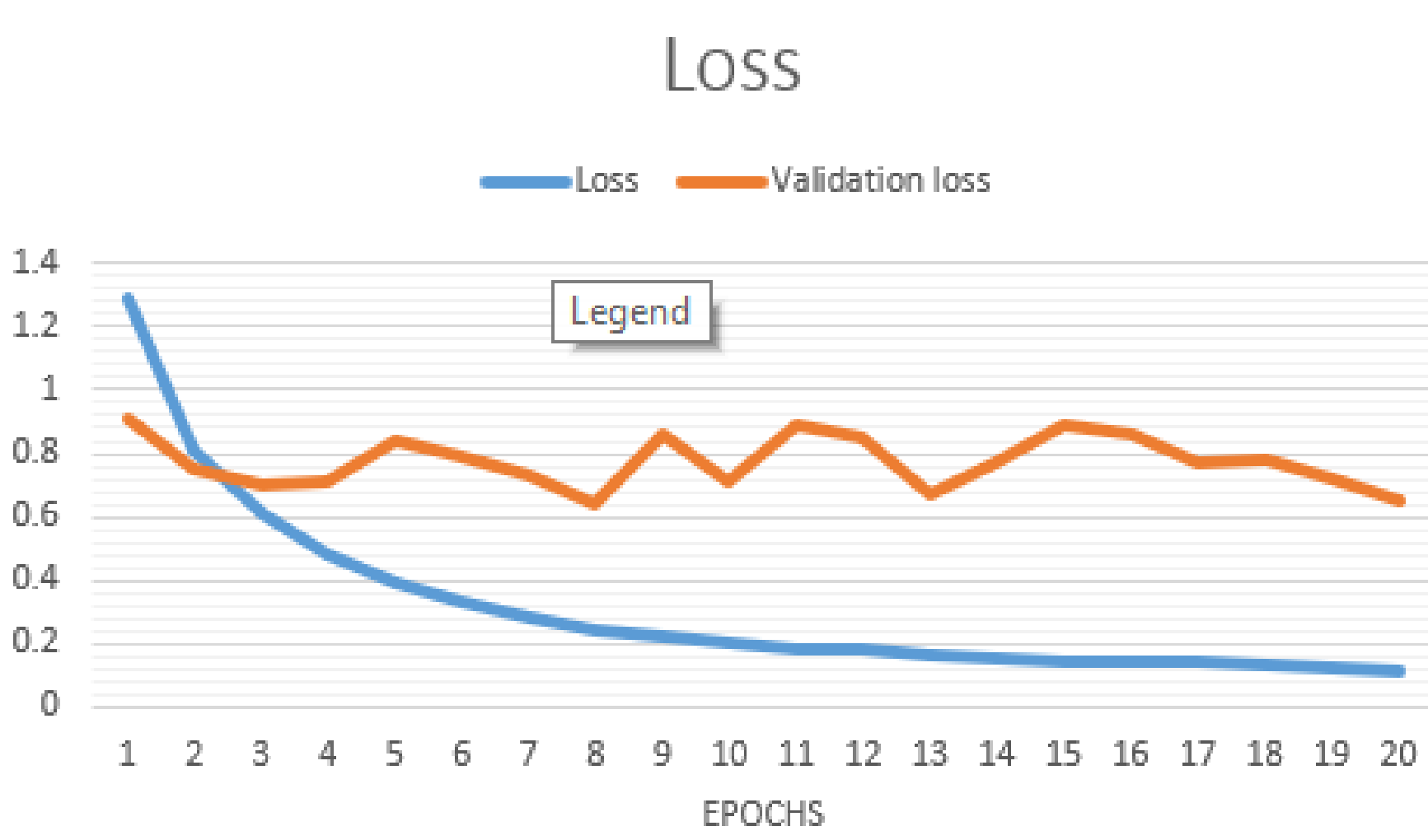
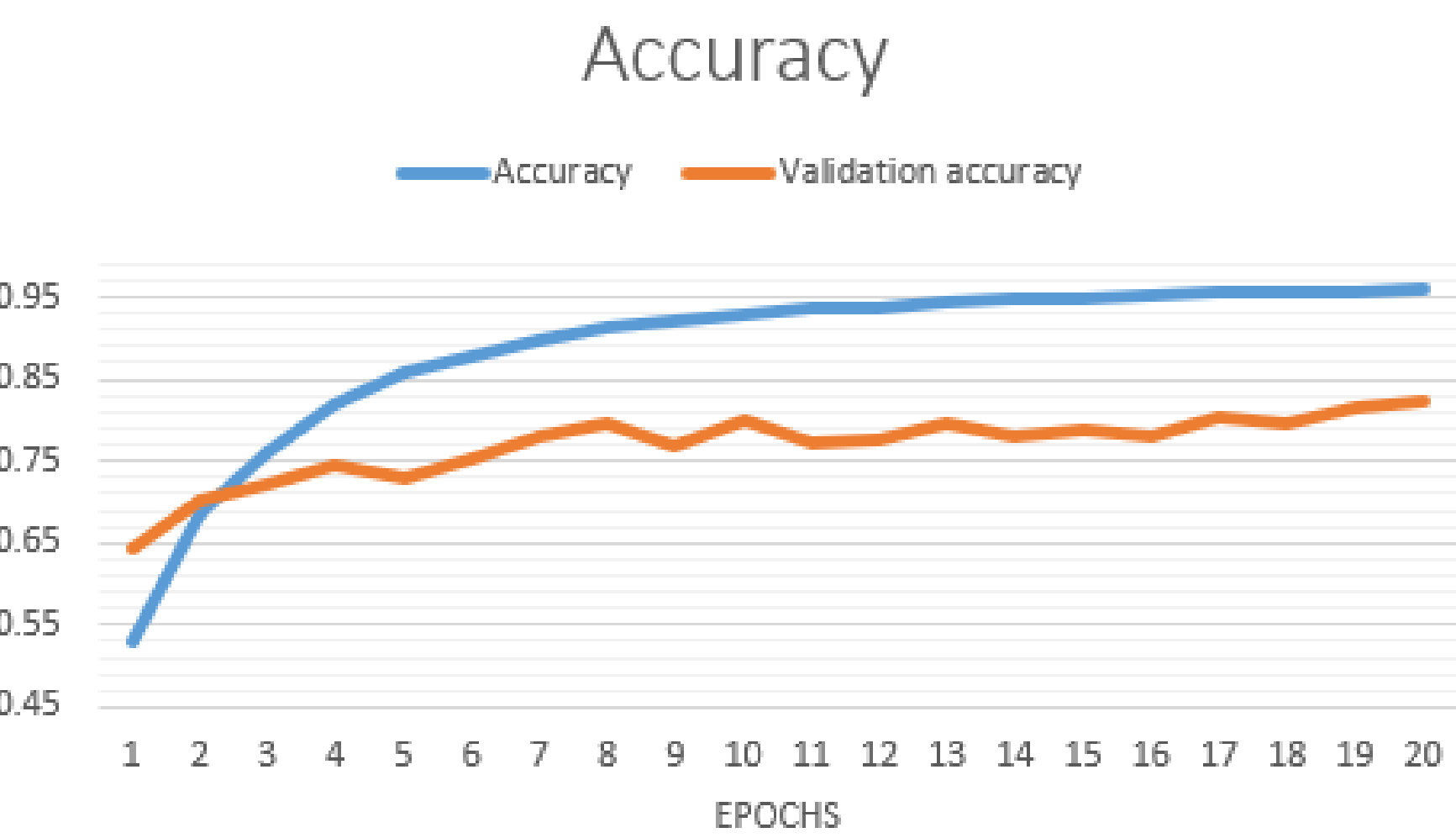
Primeri generisanih slika za detekciju sata



Primeri generisanih slika za detekciju vremena

Rezultati

Obe mreže su trenirane nad generisanim skupom podataka od 30 000 slika satova, od kojih je 20% izdvojeno za validaciju. Na datim graphicima prikazana je zavisnost *accuracy*-ja i *loss*-a od broja epoha pri treniranju neuronske mreže za detekciju vremena na satu.



Trenirani modeli su testirani nad 3 000 elemenata. U tabelama ispod dati su rezultati testiranja.

	Clock tracking
95% accurate	87.09 %
	Time reading
Hour	92.03 %
Minute	38.33 %
Minute (tolerance < 5 min)	83.24 %

Zaključak

Ovim pristupom postignuti su zadovoljavajući rezultati, ali svakako postoji prostor za poboljšanje. Jedan od načina kako bi dobili robustnije modele jeste proširenje skupa podataka novim tipovima satova i dodatnom augmentacijom podataka, kako bi model mogao još bolje da generalizuje stečeno znanje. Upotrebom algoritama kao što je ROLO, detekcija i praćenje sata bi bilo daleko performantnije, ali bi samim time postavilo i veće hardverske zahteve. Takođe, promena arhitekture mreže za prepoznavanje vremena bi potencijalno donela još bolje preformanse u kombinaciji sa obogaćenim skupom podataka.