

MODELIRANJE DINAMIKE PEČENJA ZA DOLOČANJE STANJA PEČENJA Z MREŽAMI LSTM

NINA SEDEJ,^{1,2} GREGOR KOPOREC,² JANEZ PERŠ¹

¹ Univerza v Ljubljani, Fakulteta za elektrotehniko, Ljubljana, Slovenija
nina.sedej@gorenje.com, janez.pers@fe.uni-lj.si

² Gorenje d.o.o, Ljubljana, Slovenija
nina.sedej@gorenje.com, gregor.koporec@gorenje.com

Sinopsis Osnovni način, da dosežemo dobre rezultate pri pečenju je prilagoditev časa – če želi ljubiteljski kuhar bolj zapečene piškote bo podaljšal čas peke. Ta pristop ne zagotavlja vedno istih rezultatov in lahko vodi v preveč ali premalo zapečene jedi. Za reševanje tega problema je bilo razvitih že več sistemov računalniškega vida, ni pa še bilo izvedene sistematične študije, ki bi razviti sistem primerjala z izkušenim domačim kuharjem. V tem delu predstavimo sistem računalniškega vida, ki je sestavljen iz pečiče s kamero, sistema za zajemanje slik in globokih nevronskega modelov. Delovanje sistema primerjamo z modelom ljubiteljskega kuharja. Ker se videz jedi v pečiči spreminja skozi čas, poleg konvolucijskega modela CNN uporabimo dve vrsti modelov, ki na vhodu sprejmeta zaporedje slik - CNN-LSTM in ConvLSTM. Rezultati kažejo, da model ConvLSTM prekaša model ljubiteljskega kuharja za 5 odstotnih točk v metriki F1. Da so modeli primerni za spremljanje kvalitete jedi v pečiči, morajo imeti sposobnost učenja dinamike pečenja.

Ključne besede:
dinamika pečenja,
stanje pečenja,
CNN,
CNN-LSTM,
ConvLSTM,
razvrščanje

MODELING BAKING DYNAMICS FOR DETERMINING BAKING STATE USING LSTM NETWORKS

NINA SEDEJ,^{1,2} GREGOR KOPOREC,² JANEZ PERŠ¹

¹ University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenia
nina.sedej@gorenje.com, janez.pers@fe.uni-lj.si

² Gorenje d.o.o, Ljubljana, Slovenia
nina.sedej@gorenje.com, gregor.koporec@gorenje.com

Abstract The conventional approach of getting a satisfactory baking result is to adapt baking time. If the home cook prefers the cookies to be darker, they will prolong the baking time. This approach doesn't guarantee the same results every time and can result in an underbaked or burnt dish. To solve this issue computer vision-based systems have already been developed, but there isn't any systematic study that would compare a developed computer vision system to an experienced home cook. We present a computer vision system consisting of an oven with a camera, an image acquisition system and deep learning models. We compare it to a home cook model. To model baking dynamics we also use models, that take a sequence of images as their input: CNN-LSTM and ConvLSTM. We conclude that ConvLSTM outperforms the home cook model by 5 percentage points in F1 metric. A crucial attribute for machine learning models to be successfully used for monitoring baking quality is learning baking dynamics.

Keywords:

baking dynamics,
baking state,
CNN,
CNN-LSTM,
ConvLSTM,
classification

1 Uvod

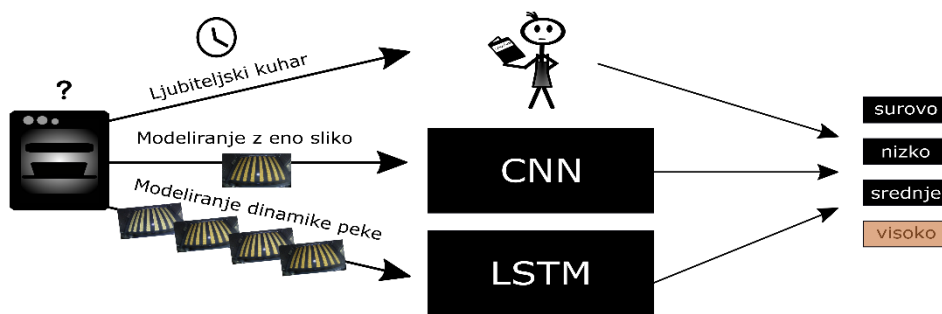
Peka jedi je kompleksen nelinearen proces (Pylar in Gorton, 2008), ki ga lahko v splošnem razdelimo na dve fazi. V prvi fazi je dinamika procesa počasna, saj ne opazimo vidnih sprememb. Sledi faza hitre dinamike, kjer zaradi kompleksnih kemičnih procesov živilo poleg okusa spremeni tudi svoj videz (obliko, barvo in teksturo) (Purlis, 2010). Videz tako postane bistven dejavnik za modeliranje stanja pečenja (Purlis, 2010). Izraz stanje pečenja pomeni oceno gotovosti jedi (surova, nizko, srednje ali visoko zapečena).

Ugotavljanje stanja pečenja v dinamični fazi peke je pomemben vidik poklicnega kuharja, saj se poklicni kuhar lahko nauči, kako čas pečenja vpliva na kakovost gotove jedi. Čas pečenja tako postane bistven parameter recepta, s katerim poenostavimo zapletenost postopka peke. Recept lahko uporabi ljubiteljski kuhar, pri čemer lahko s spremembo časa pečenja jed prilagodi svojim željam. Na primer, če ima rajši temnejše piškote, jih lahko peče nekaj minut dlje, kot piše v receptu.

Prilagoditev časa peke za zagotavljanje zelenih rezultatov ima svoje pomanjkljivosti: (i) Zgolj sledenje receptu nam ne zagotavlja enakih rezultatov, saj je postopek peke odvisen tudi od drugih dejavnikov (npr. vrsta pečice, temperatura, vrsta pladnja, variacija v sestavinah v različnih državah, razlika v masi sestavin, ipd.). Jed, ki je pripravljena po enakem receptu, je zaradi tovrstnih dejavnikov lahko gotova nekaj minut prej ali pozneje. (ii) Dinamična faza lahko ob koncu pečenja povzroči hitre spremembe videza jedi. Ljubiteljski kuharji tako lahko hitro podcenijo dinamiko pečenja, rezultat pa je preveč pečena ali celo zažgana jed. V tem primeru jed ni samo neužitna, ampak tudi nevarna za zdravje (Rudén, 2004). (iii) Ko kuharji odpirajo pečico, da ugotovijo, ali je jed gotova, povzročijo padec temperature, kar zmoti proces pečenja.

Za reševanje tovrstnih problemov so bile zato razvite podporne tehnologije na osnovi računalniškega vida. Običajno delujejo tako, da sistem računalniškega vida zajame vizualno informacijo, iz katere s tehnikami stojnega učenja razberemo spremembe stanja jedi. Vendar pa do sedaj še ni bilo izvedene študije, kjer bi sistematsko primerjali razviti sistem računalniškega vida, z metodo prilagajanja časa, ki jo uporabljajo ljubiteljski kuharji. Tako ni jasno, *ali je sistem računalniškega vida sploh primerna rešitev za določanje stanja jedi v pečici?*

Da bi odgovorili na to vprašanje, v tem delu sistematično primerjamo trenutno najboljši sistem računalniškega vida z metodo prilagajanja časa pečenja, ki ga uporabljajo domači kuharji – *model ljubiteljskega kuharja*. Slika 1 prikazuje shematični prikaz prikazanega pristopa. V ta namen smo razvili sistem računalniškega vida, ki je sestavljen iz pečice s kamero in prilagojeno osvetlitvijo ter modela globoke nevronske mreže za razpoznavanje stanja pečenja: je jed surova ali nizko, srednje ali visoko zapečena. Model ljubiteljskega kuharja simulira odločitve izkušenega ljubiteljskega kuharja in predstavlja zahtevno osnovo za primerjalno analizo modelov globokih nevronskih mrež.



Slika 1: Shematični prikaz predlaganega pristopa. Primerjamo tri modele za spremljanje stanja pečenja: ljubiteljskega kuharja, ki peče po receptu, CNN, ki procesira po eno sliko naenkrat in dve LSTM arhitekturi, ki procesirata časovno sosledje več slik.

Vir: lasten.

Ker se poklicni kuharji učijo kompleksne dinamike pečenja z opazovanjem pečenja skozi čas, pričakujemo, da je mogoče učinkovito modelirati dinamiko pečenja z modeli, ki izkoriščajo časovno sosledje podatkov. Domnevamo, da bodo arhitekture LSTM, ki na vходу sprejmejo časovno zaporedje slik, izboljšale rezultate modelov CNN, ki delujejo na posamičnih slikah. Da bi preverili to predpostavko, v tem delu analiziramo dve različni arhitekturi LSTM: CNN-LSTM in ConvLSTM.

Rezultati študije kažejo, da modeli CNN niso primerni za razpoznavanje stanja jedi, saj so slabši od modela ljubiteljskega kuharja. Arhitekture ConvLSTM so po drugi strani obetavne, saj ljubiteljskega kuharja prekašajo za 5 odstotnih točk v metriki F1. Rezultati potrjujejo domnevo: *Ključna značilnost modelov za spremljanje stanja peke je sposobnost učenja dinamike pečenja.*

2 Sorodna dela

Kakovost izdelkov, ki se pečejo v pečici lahko določamo iz meritev različnih senzorjev, kot so temperaturna sonda, spektrometer ali kolorimeter (Ureta et al., 2014; Ureta et al., 2016). Težava pri tovrstnih analizah je, da so destruktivne. Kadar želimo določiti lastnosti jedi, jo moramo vzeti iz pečice, to pa vpliva na proces pečenja. V izogib težavam kontaktnega merjenja, je bilo predlaganih več neinvazivnih metod računalniškega vida (Mogol in Gökmen, 2014).

Sistem računalniškega vida lahko uporablja vizualne lastnosti kot sta barva in oblika hrane za oceno njenih lastnosti. Shafiee et al. (2014) so uporabili sistem računalniškega vida za določanje barve medu in uporabili preprost model nevronske mreže za napovedovanje lastnosti medu in določitev njegove kakovosti. Purlis in Salvadori (2009) sta predlagala kolorimetrično metodo za modeliranje in napovedovanje sprememb v barvi kruhove skorje med peko s pomočjo matematičnega modela za kinetično rjavenje. Paquet-Durand et al. (2012) so pri spremljanju procesa peke žemljic uporabili tudi informacijo o spremembi velikosti.

Konvolucijske nevronske mreže (angl. Convolutional Neural Networks – CNN) so zelo uspešne pri reševanju številnih nalog, povezanih s slikami (Krizhevsky et al., 2012; He et al., 2016). Cotrim et al. (2020) so zato predlagali kolorimetrični pristop, v katerem s pomočjo CNN razvrščajo barvo kruhove skorje v sedem različnih stanj. Kljub temu do sedaj še ni bilo izvedene sistematične študije, kjer bi primerjali modele globokih nevronskih mrež s sposobnostjo ljubiteljskega kuharja, da s prilagoditvijo časa pečenja dobi optimalne rezultate.

Razredi modelov, ki so pokazali dobro delovanje s časovno soslednimi podatki, so mreže z dolgim kratkoročnim spominom (angl. Long Short-Term Memory - LSTM) (Hochreiter in Schmidhuber, 1997). Za uporabo mrež LSTM z zaporedji slik je bilo predlaganih več rešitev, med katerimi izstopata različici CNN-LSTM in ConvLSTM. Različica CNN-LSTM je sestavljena iz komponente CNN, ki jo uporabimo kot določevalnik značilk, in LSTM komponente, od koder dobimo končne rezultate razpoznavalnika (Donahue et al., 2017). Različica ConvLSTM je arhitekturno bolj podobna običajni LSTM celici, le da uteži nadomestijo konvolucijska jedra, kar omogoča pretok podatkov skozi časovno mrežo v izvorni 3D obliki (Shi et al., 2015).

Arhitekture CNN-LSTM so bile uporabljene pri različnih opravilih, kot je opisovanje (Gao et al., 2017; Donahue et al., 2017) in razvrščanje video posnetkov (Ng et al., 2015). Po drugi strani so (Yadav et al. 2022) model ConvLSTM uporabili za razpoznavanje človeških kretenj na podlagi posnetkov vzorcev gibanja. (Shi et al. 2015) so model ConvLSTM uporabili za napovedovanje padavin na podlagi prostorsko-časovnih radarskih podatkov. V okviru obeh raziskav so ugotovili, da je model ConvLSTM dosegel boljše rezultate kot model CNN-LSTM. Nismo zasledili raziskav, v katerih bi bila katera od različic mreže LSTM uporabljena za spremljanje kakovosti hrane pri pečenju v pečici.

3 Podatkovna zbirka

3.1 Sistem za zajem podatkov

Sistem za zajem podatkov je bil sestavljen iz prototipa pečice, digitalne kamere in računalnika. Prototip pečice s kamero je prispevalo podjetje Gorenje d.o.o. Kamera je bila nameščena v zgornji del ohišja pečice in je zajemala slike skozi odprtino, ki je bila izolirana s steklom, odpornim na visoke temperature. Za izboljšanje kakovosti slike smo obstoječo osvetlitev v pečici dopolnili z dvema LED trakovima. Uporabili smo 1/2,7-palčno RGB kamero, ki zagotavlja ločljivost 1920×1080 slikovnih pik. S kamero smo vsakih 30 sekund zajemali fotografije plitkega emajliranega pekača dimenzije 460×360 mm, na katerem se je pekla jed. Ker je delovna temperatura kamere 105 °C, temperature v pečici pa se lahko gibljejo tudi do 400 °C (piroliza), smo za hlajenje kamere izdelali ventilatorski sistem. Kamera je bila povezana na računalnik Raspberry Pi 4, ki je nadzoroval zajemanje slik s programsko opremo Motion (motion-project, 2020) različice 4.3.2.

3.2 Peka piškotov

Pekli smo brizgane piškote, ki so bili pripravljene po standardnem receptu za določanje enakomernosti porazdelitve toplote na bremenu. Izraz pečenje označuje celoten postopek peke od postavitve pladnja s surovo jedjo v pečico do trenutka, ko jed doseže zlatorjavo bravo (kategorija visoko zapečeno) in jo vzamemo iz pečice. Posamezna peka je trajala okvirno 33 min. S posamezno peko smo pridobili v povprečju 60 slik. Da bi bili pridobljeni podatki čim bolj raznovrstni, smo pekli piškote na 35 različnih načinov. Ti vključujejo tri različne programe peke (predgretje,

klasična peka in vroči zrak), dve različni temperaturi (140 °C in 145 °C), pekača dveh različnih barv (siva in črna) ter dva nivoja v pečici (srednji in spodnji nivo). Poleg tega smo uporabili dve različni razporeditvi piškotov na pekaču (pravokotna in vzporedna) ter različno količino piškotov (od 5 do 11 vrstic).

Na koncu vsakega poskusa smo piškote prestavili iz pečice v testno komoro, kjer smo s spektrofotometrom izmerili njihovo barvo in določili končno stanje jedi. Na podlagi meritve končnega stanja smo nato določili kategorije slik (glej razdelek 3.4 za podrobnosti).

3.4 Podatkovna zbirka OvenRecognition

Podatkovna zbirka OvenRecognition je sestavljena iz 32990 slik brizganih piškotov. Primeri vzorcev iz podatkovne zbirke so prikazani na sliki 2. 28230 slik je namenjenih učenju modelov, 4760 slik sodi v testno množico. Vsak vzorec vsebuje podatke o času in programu peke, ID oznako pečenja ter stanje zapečenosti. Stanje zapečenosti določajo štiri kategorije: surovo, nizko zapečeno, srednje zapečeno in visoko zapečeno. V nasprotju s kategorijo surovo, ostale kategorije predstavljajo užitne vzorce. Podatkovna zbirka je glede na stanje zapečenosti močno neuravnotežena (npr. 22970 vzorcev spada v kategorijo surovo), kar je posledica dinamike procesa pečenja.



Slika 2: Primeri slik iz zbirke podatkov OvenRecognition, ki predstavljajo vzorce iz kategorij surovo, nizko zapečeno, srednje zapečeno in visoko zapečeno.

Vir: lasten.

4 Modeliranje procesa pečenja

4.1 Model ljubiteljskega kuharja

Kot osnovo za primerjavo modelov smo uporabili model ljubiteljskega kuharja, ki razvršča slike na podlagi časa, ki je preteklo od začetka pečenja. Namen te metode je simulirati uporabnika pečice, ki upošteva čas peke, kot je naveden v receptu. Časovne

meje med kategorijami smo izbrali na način, da zmanjšamo število napačno razvrščenih slik na učni množici podatkovne zbirke. Na ta način smo simulirali proces odločanja izkušenega ljubiteljskega kuharja, in jih hkrati naredili bolj nepristranske. Meje smo postavili na naslednje vrednosti: 0–25 min za surovo, 25,5–28 min za nizko zapečeno, 28,5–33 min za srednje zapečeno ter 33,5 min in več za visoko zapečeno kategorijo.

4.2 CNN

Za model CNN smo izbrali EfficientNet_b4, ki je bil prednaučen na podatkovni zbirki ImageNet (Deng et al., 2009). Model prihaja iz družine razširljivih modelov, ki z relativno malo hiperparametri dosegajo dobre rezultate (Tan in Le, 2019). Zadnji, polno povezan sloj smo doučili na podatkovni zbirki OvenRecognition. Začetna učna stopnja je bila 10^{-2} . Učno stopnjo smo zmanjšali za faktor 10 vsakič, ko se izguba ni izboljšala v 20 zaporednih epochah učenja. Učne podatke smo na začetku vsake epohe premešali. Neuravnoteženost podatkovne zbirke smo rešili z uteževanjem kriterijske funkcije navzkrižne entropijske izgube (angl. Cross Entropy Loss). Pri tem smo uporabili inverzno relativno frekvenco razreda.

4.3 CNN-LSTM

Uporabili smo dva različna modela CNN-LSTM. V prvem modelu, poimenovanem AlexCNN-LSTM smo za luščenje značilik uporabili model AlexNet (Krizhevsky et al., 2012). V drugem modelu, EfficientCNN-LSTM, smo za luščenje značilik uporabili model EfficientNet_b4 (Tan in Le, 2019). Obe CNN komponenti sta bili prednaučeni na podatkovni zbirki ImageNet (Deng et al., 2009). Za vhodno zaporedje slik smo uporabljali 4 časovno sosedne slike. LSTM komponente modelov smo doučili na podatkovni zbirki OvenRecognition. Doučevanje je potekalo na enak način kot pri modelu CNN, le da je bila začetna učna stopnja 10^{-3} .

4.4 ConvLSTM

Uporabili smo model ConvLSTM iz (Shi et al., 2015), ki smo ga prilagodili tako, da je sestavljen iz začetnega konvolucijskega sloja, ki mu sledi sloj ConvLSTM. Naučili smo dva modela ConvLSTM z različno velikostjo vhodnega zaporedja slik. ConvLSTM_n2 predstavlja model, ki uporablja 2 časovno sosedni sliki za razpoznavanje. Z ConvLSTM_n4 smo označili model, ki uporablja 4 časovno

sosledne slike. Modela ConvLSTM smo naučili izključno na podatkovni zbirki OvenRecognition. Učenje je potekalo na enak način kot pri modelu CNN, le da je bila začetna učna stopnja 10^{-4} .

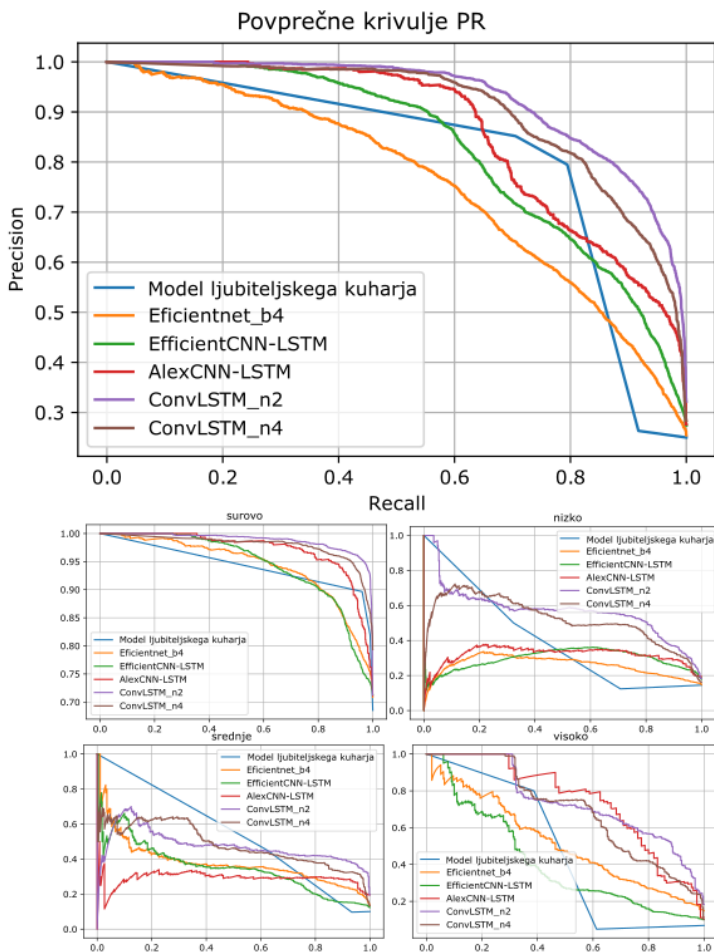
5 Eksperimenti

Da bi ocenili, ali je uporaba globokih nevronskih mrež primerna rešitev za določevanje stanja jedi v pečici, smo najprej primerjali izbran model CNN z modelom ljubiteljskega kuharja. Pričakovali smo slabe rezultate, saj menimo, da modeli CNN po naravi ne zmorejo natančneje modelirati dinamike pečenja, ki jo lahko zasledimo z opazovanjem skozi čas. Po drugi strani smo pričakovali, da se dinamike pečenja lahko naučijo modeli CNN-LSTM, saj razpoznavanje na podlagi časovnega zaporedja slik. Pri tovrstnih modelih smo dodatno preverili, kako različna kapaciteta določevalcev značilnik vpliva na rezultate. Ker sorodna dela nakazujejo na dodatne izboljšave pri uporabi modelov ConvLSTM, smo preverili tudi njihovo delovanje pri razpoznavanju stanja pečenja. Pri eksperimentih s tovrstnimi modeli smo preverili še vpliv dolžine časovnega zaporedja slik.

Za evaluacijo modelov smo uporabili pet različnih metrik: Precision, Recall, F1, krivuljo PR in normalizirano matriko zamenjav (angl. confusion matrix).

6 Rezultati in diskusija

Slika 3 prikazuje povprečne krivulje PR in krivulje PR za vsak razred posebej. Rezultati povprečnih krivulj PR nakazujejo, da model Efficientnet_b4 ni zmožen dosegati rezultatov ljubiteljskega kuharja. S podrobno analizo krivulj PR za posamezne razrede smo ugotovili, da Efficientnet_b4 za surovo jed prekaša ljubiteljskega kuharja do vrednosti Recall 0,8. Menimo, da bi to lahko bila posledica neuravnoteže podatkovne zbirke, saj je surova kategorija močno zastopana. Delovanje modela Efficientnet_b4 za ostale kategorije je bistveno slabše od ljubiteljskega kuharja. Rezultati, prikazani v tabeli 1, prav tako nakazujejo na slabo delovanje modela Efficientnet_b4. Razlika med modelom Efficientnet_b4 in ljubiteljskim kuharjem v metriki F1 je 11,34 odstotnih točk. To potrjuje naša pričakovanja in nakazuje, da je učenje dinamike peke ključna značilnost za uspešno spremljanje stanja pečenja.



Slika 3: Krivulje PR za surovo, nizko, srednje in visoko zapečeno kategorijo. Večja površina pod krivuljo v splošnem pomeni boljše delovanje. Modela ConvLSTM dosejata boljše rezultate za večino vrednosti Recall-a. Modeli CNN in LSTM ne dosežajo boljših rezultatov od modela ljubiteljskega kuharja.

Vir: lasten.

Tabela 1: Vrednosti metrik Precision (P), Recall (R) in F1 za posamezne modele. Modela ConvLSTM dosežeta najboljše rezultate v vseh metrikah.

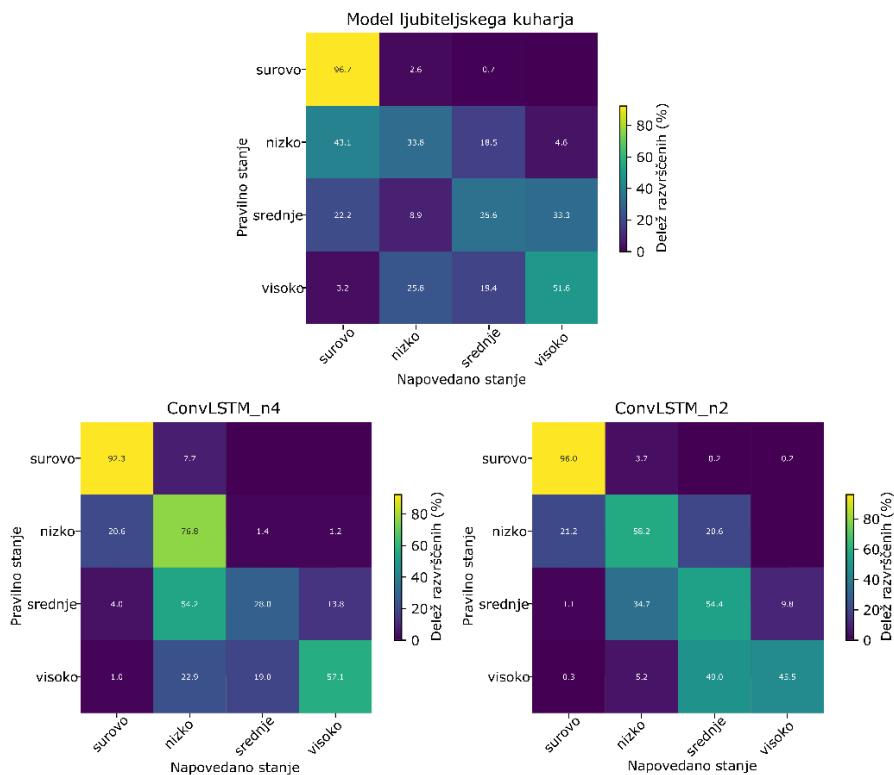
Model	P	R	F1
Model ljubiteljskega kuharja	66,21	57,33	61,45
EfficientNet_b4	47,73	52,73	50,11
AlexCNN-LSTM	63,27	39,16	48,37
EfficientCNN-LSTM	55,61	41,53	47,56
ConvLSTM_n2	68,07	63,52	65,72
ConvLSTM_n4	69,72	63,54	66,49

vir: lasten

Pri analizi rezultatov modelov CNN-LSTM smo ugotovili, da tovrstni modeli prekašajo model CNN EfficientNet_b4 po krivuljah PR in metriki Precision, metriki Recall in F1 pa nakazujeta slabše delovanje. Opazimo lahko, da je natančnost modelov CNN-LSTM slabša od ljubiteljskega kuharja na območju Recall od 0,6 do 0,8. Pri primerjavi med različnimi kapacitetami določevalnikov značilk v modelih CNN-LSTM smo ugotovili, da povečana kapaciteta daje slabše rezultate. Rezultati bi lahko nakazovali, da se CNN-LSTM modeli do določene mere učijo dinamike pečenja, vendar ne dovolj. Verjetno ima velik vpliv komponent CNN modelov, ki ne zmore luščiti značilk z vsebovano dinamiko pečenja. Ločevanje značilk na posamezne razrede v nadaljnjih slojih modela pa je posledično oteženo. Vpliv komponent CNN se lahko veča z večanjem kapacitete določevalnika, kar bi bil lahko vzrok slabšega delovanja modela EfficientCNN-LSTM.

Po drugi strani modeli ConvLSTM presegajo rezultate modela ljubiteljskega kuharja, modelov CNN-LSTM in modela CNN EfficientNet_b4 po krivuljah PR. Boljše delovanje nakazujejo tudi ostale metrike. Rezultati med ConvLSTM_n4 in modelom ljubiteljskega kuharja se razlikujejo za 3,51 odstotne točke v metriki Precision, 6,21 odstotnih točk v metriki Recall in 5,01 odstotnih točk v metriki F1. Za natančnejšo primerjavo med modelom ljubiteljskega kuharja in modeli ConvLSTM smo primerjali njihove matrike zamenjav, ki so prikazane na sliki 4. Matrike so izraziteje diagonalne za modele ConvLSTM, kar pomeni, da je bilo pravilno razvrščenih več vzorcev. Še več, modeli ConvLSTM razvrstijo bistveno manj užitnih vzorcev (nizko, srednje in visoko zapečeno) med surove kot model ljubiteljskega kuharja.

Da bi pridobili dodaten vpogled v delovanje modela ConvLSTM smo preverili še vpliv dolžine časovnega zaporedja slik. Model ConvLSTM_n4 je dosegel boljše rezultate kot ConvLSTM_n2 v vseh metrikah (tabela 1), vendar razlike niso statistično značilne. Sklepamo lahko, da zmanjšanje število vhodnih slik ni bistveno vplivalo na sposobnost modeliranja dinamike pečenja modela ConvLSTM. Iz slike 4 je razvidno, da je ConvLSTM_n4 boljši pri razvrščanju kategorije nizke in visoke zapečenosti, vendar pa označi več surovih vzorcev kot užitne, kar pa ni zaželeno. Model s krajšim časovnim zaporedjem slik zato predstavlja boljšo izbiro za spremljanje stanja jedi med pečenjem v pečici.



Slika 4: Primerjava matrik zamenjav za model ljubiteljskega kuharja in modela ConvLSTM_n4 in ConvLSTM_n2.

Vir: lasten.

7 Zaključek

Osnovni pristop za doseganje zadovoljivega rezultata pečenja v pečici je prilagajanje časa pečenja. Ta pristop ne zagotavlja vedno enakih rezultatov in nemalokrat povzroči zažgano jed. Da bi rešili to težavo smo razvili sistem računalniškega vida s katerim brezkontaktno zajemamo slike jedi v notranjosti pečice. S pomočjo modela globokega učenja razpoznavamo stanje jedi v pečici. Razviti sistem računalniškega vida sistematično primerjamo z modelom ljubiteljskega kuharja, ki simulira odločitve izkušenega uporabnika pečice.

Poklicni kuharji se učijo kompleksne dinamike učenja z opazovanjem pečenja jedi skozi čas, zato pričakujemo, da bodo arhitekture LSTM, ki izkoriščajo časovno sosledje podatkov, izboljšale rezultate modela CNN. Da bi preizkusili to hipotezo

smo poleg modela CNN naučili in analizirali tudi štiri modele arhitektur CNN-LSTM in ConvLSTM. Rezultati študije kažejo, da modeli CNN in CNN-LSTM niso primerni za razpoznavanje stanja pečenja. Arhitekture ConvLSTM so po drugi strani dosegle obetavne rezultate, saj med drugim v metriki F1 prekašajo model izkušenega ljubiteljskega kuharja za 5 odstotnih točk. Končna ugotovitev te raziskave je, da je učenje dinamike peke ključna značilnost za uspešno uporabo modelov strojnega učenja za spremljanje stanja jedi med pečenjem v pečici.

Ta ugotovitev odpira nove raziskovalne priložnosti, s katerimi bi lahko dodatno izboljšali spremljanje kakovosti peke. Ker se modeli LSTM osredotočajo le na nekaj zaporednih slik, bi lahko razvili arhitekture, ki bi spremljale celoten proces pečenja.

Opombe

Raziskavo je omogočilo podjetje Gorenje d.o.o.

Literatura

- Cotrim, W. d., Minim, V. P., Felix, L. B., & Minim, L. A. (2020). Short convolutional neural networks applied to the recognition of the browning stages of bread crust. *Journal of Food Engineering*, 277, 109916. doi:10.1016/j.jfoodeng.2020.109916
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, (str. 248–255). doi:10.1109/CVPR.2009.5206848
- Donahue, J., Hendricks, L. A., Rohrbach, M., Venugopalan, S., Guadarrama, S., Saenko, K., & Darrell, T. (2017). Long-Term Recurrent Convolutional Networks for Visual Recognition and Description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 677–691. doi:10.1109/TPAMI.2016.2599174
- Gao, L., Guo, Z., Zhang, H., Xu, X., & Shen, H. T. (2017). Video Captioning With Attention-Based LSTM and Semantic Consistency. *IEEE Transactions on Multimedia*, 19, 2045–2055. doi:10.1109/TMM.2017.2729019
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (str. 770–778). doi:10.1109/CVPR.2016.90
- Hochreiter, S., & Schmidhuber, J. (December 1997). Long Short-term Memory. *Neural computation*, 9, 1735–80. doi:10.1162/neco.1997.9.8.1735
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012, January). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*, 25. doi:10.1145/3065386
- Meenu, M., Kurade, C., Neelapu, B. C., Kalra, S., Ramaswamy, H. S., & Yu, Y. (2021). A concise review on food quality assessment using digital image processing. *Trends in Food Science & Technology*, 118, 106–124. doi:10.1016/j.tifs.2021.09.014
- Mogol, B. A., & Gökmen, V. (May 2014). Computer vision-based analysis of foods: A non-destructive colour measurement tool to monitor quality and safety. *J. Sci. Food Agric.*, 94, 1259–1263. doi:10.1002/jsfa.6500

- motion-project. (2020). Motion project. *Motion project*. Pridobljeno iz <https://motion-project.github.io/>
- Ng, J., Hausknecht, M., Vijayanarasimhan, S., Vinyals, O., Monga, R., & Toderici, G. (June 2015). *Beyond short snippets: Deep networks for video classification*. doi:10.1109/CVPR.2015.7299101
- O'Shea, K., & Nash, R. (November 2015). An Introduction to Convolutional Neural Networks. *ArXiv e-prints*.
- Paquet-Durand, O., Solle, D., Schirmer, M., Becker, T., & Hitzmann, B. (2012). Monitoring baking processes of bread rolls by digital image analysis. *Journal of Food Engineering*, 111, 425–431. doi:10.1016/j.jfoodeng.2012.01.024
- Purlis, E. (2010). Browning development in bakery products - A review. *Journal of Food Engineering*, 99, 239–249. doi:10.1016/j.jfoodeng.2010.03.008
- Purlis, E., & Salvadori, V. O. (2009). Modelling the browning of bread during baking. *Food Research International*, 42, 865–870. doi:10.1016/j.foodres.2009.03.007
- Pylar, E. J., & Gorton, L. A. (2008). *Baking science & technology: volume I: fundamentals & ingredients*. Sosland Pub.
- Rudén, C. (2004). Acrylamide and cancer risk—expert risk assessments and the public debate. *Food and Chemical Toxicology*, 42, 335–349. doi:10.1016/j.fct.2003.10.017
- Shafiee, S., Minaei, S., Moghaddam-Charkari, N., & Barzegar, M. (2014). Honey characterization using computer vision system and artificial neural networks. *Food Chemistry*, 159, 143–150. doi:10.1016/j.foodchem.2014.02.136
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W. K., & Wang-chun, W. O. (June 2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting.
- Tan, M., & Le, Q. (May 2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*.
- Ureta, M. M., Olivera, D. F., & Salvadori, V. O. (2014). Baking of muffins: Kinetics of crust color development and optimal baking time. *Food and Bioprocess Technology*, 7, 3208–3216. doi:10.1007/s11947-014-1292-z
- Ureta, M., Olivera, D., & Salvadori, V. (August 2016). Influence of baking conditions on the quality attributes of sponge cake. *Food Science and Technology International*, 23. doi:10.1177/1082013216666618
- Yadav, S. K., Tiwari, K., Pandey, H. M., & Akbar, S. A. (2022). Skeleton-based human activity recognition using ConvLSTM and guided feature learning. *Soft Computing*, 26, 877–890. doi:10.1007/s00500-021-06238-7
- Zhu, L., Spachos, P., Pensini, E., & Plataniotis, K. N. (2021). Deep learning and machine vision for food processing: A survey. *Current Research in Food Science*, 4, 233–249. doi:10.1016/j.crfs.2021.03.009