

# Evaluacija modela za estimaciju kalorija hrane prikupljanjem standardiziranog skupa podataka

---

**Burazin, Marko**

**Master's thesis / Diplomski rad**

**2024**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Split, Faculty of Science / Sveučilište u Splitu, Prirodoslovno-matematički fakultet**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:166:710317>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2025-02-20**

*Repository / Repozitorij:*

[Repository of Faculty of Science](#)



UNIVERSITY OF SPLIT



DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

SVEUČILIŠTE U SPLITU  
PRIRODOSLOVNO MATEMATIČKI FAKULTET

DIPLOMSKI RAD

**EVALUACIJA MODELA ZA ESTIMACIJU  
KALORIJA HRANE PRIKUPLJANJEM  
STANDARDIZIRANOG SKUPA PODATAKA**

Marko Burazin

Split, rujan 2024.

# Temeljna dokumentacijska kartica

Diplomski rad

Sveučilište u Splitu  
Prirodoslovno-matematički fakultet  
Odjel za informatiku  
Ruđera Boškovića 33, 21000 Split, Hrvatska

## EVALUACIJA MODELA ZA ESTIMACIJU KALORIJA HRANE PRIKUPLJANJEM STANDARDIZIRANOG SKUPA PODATAKA

Marko Burazin

### SAŽETAK

U ovom radu evaluiran je model za estimaciju kalorijskih vrijednosti hrane koji koristi napredne tehnike dubokog učenja. Model je treniran i testiran unutar Python okruženja u Google Colab-u, a korištene su arhitekture poput MobileNetV3 i NASNet. Ove arhitekture omogućuju simultano rješavanje problema klasifikacije i regresije što omogućuju visoku razinu točnosti u prepoznavanju hrane, sastojaka, nutritivnih vrijednosti i estimaciji kalorijskih vrijednosti. Fokus istraživanja je na prikupljanju, analizi i standardizaciji skupova podataka kako bi se osigurala dosljednost i pouzdanost podataka na kojima model trenira. Svi skupovi podataka koji su korišteni su javno dostupni na Internetu. Razvijen je cjevovod za automatizaciju obrade skupova podataka, čime se značajno ubrzava i pojednostavljuje rad s velikim količinama podataka. Rezultati eksperimenta pokazuju napredak u preciznosti estimacije kalorija korištenjem zajedničkog skupa podataka, čime se potvrđuje potreba za izradom velikog skupa podataka te njegov doprinos unaprjeđenju tehnologija u području procjene nutritivnih vrijednosti hrane.

**Ključne riječi:** estimacija kalorija, duboko učenje, strojno učenje, MobileNetV3, NASNet, standardizacija podataka, cjevovod

Rad je pohranjen u knjižnici Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

**Rad sadrži:** 57 stranica, 26 grafičkih prikaza, 14 tablica i 56 literaturnih navoda. Izvornik je na hrvatskom jeziku.

**Mentor:** **Dr. sc. Saša Mladenović**, redoviti profesor Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

**Ocjenjivači:** **Dr. sc. Saša Mladenović**, redoviti profesor Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

**Nika Jerković**, asistent Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

**Antonela Prnjak**, asistent Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

Rad prihvaćen: **rujan 2024.**

# Basic documentation card

Thesis

University of Split  
Faculty of Science  
Department of Computer Science  
Ruđera Boškovića 33, 21000 Split, Croatia

## EVALUATION OF A MODEL FOR FOOD CALORIE ESTIMATION BY COLLECTING A STANDARDIZED DATASET

Marko Burazin

### ABSTRACT

In this paper, a model for estimating caloric values of food that uses advanced deep learning techniques was evaluated. The model was trained and tested within the Python environment at Google Colab, and architectures such as MobileNetV3 and NASNet were used. These architectures enable simultaneous solving of classification and regression problems, which enable a high level of accuracy in recognizing food, ingredients, nutritional values and caloric value estimation. The focus of the research is on the collection, analysis and standardization of data sets to ensure consistency and reliability of the data on which the model is trained. All datasets used are publicly available on the Internet. A pipeline for automating the processing of data sets has been developed, which significantly speeds up and simplifies work with large amounts of data. The results of the experiment show progress in the accuracy of calorie estimation using a common data set, thus confirming the need for the creation of a large data set and its contribution to the improvement of technologies in the field of nutritional assessment of food

**Key words:** calories estimation, deep learning, machine learning, MobileNETV3, NASNet, standardization of data, data pipeline

Thesis deposited in library of Faculty of science, University of Split

**Thesis consists of:** 57 pages, 26 figures, 14 tables and 56 references

Original language: Croatian

**Mentor:** Saša Mladenović, Ph.D. *Professor of Faculty of Science, University of Split*

**Reviewers:** Saša Mladenović, Ph.D. *Professor of Faculty of Science, University of Split*

Nika Jerković, *Instructor of Faculty of Science, University of Split*

Antonela Prnjak, *Instructor of Faculty of Science, University of Split*

Thesis accepted: **September 2024.**



## Sadržaj:

<b>Uvod</b> .....	<b>1</b>
<b>1. Pregled literature</b> .....	<b>2</b>
<b>1.1. Arhitektura procesa određivanja kalorija</b> .....	<b>3</b>
1.1.2. Skupovi podataka za prepoznavanje hrane .....	4
1.1.3. Kratki pregled trenutno poznatih skupova podataka I njihovog sadržaja .....	5
1.1.4. Izazovi trenutno poznatih skupova podataka .....	6
1.1.5. Detekcija objekata pomoću dubokog učenja .....	7
1.1.6. Segmentacija slike pomoću dubokog učenja .....	8
1.1.7. Metode određivanja volumena hrane .....	9
1.1.8. Određivanje kalorija .....	10
1.1.9. Dijagram toka.....	11
<b>1.2. Korišteni Algoritmi i dosadašnji rezultati na specifičnim skupovima podataka</b> .....	<b>12</b>
<b>2. Pronalaženje, istraživanje i standardizacija skupova podataka za hranu</b> .....	<b>14</b>
<b>2.1. Izvori skupova podataka na Internetu</b> .....	<b>15</b>
<b>2.2. Dostupni skupovi podataka na Internetu i njihove osnovne značajke</b> .....	<b>16</b>
2.2.1. Pristup estimacije jedne porcije jela na slici.....	19
<b>2.3. Karakteristike i sadržaj odabranih skupova podataka</b> .....	<b>20</b>
2.3.1. FOOD 101 .....	20
2.3.2. Recipes5k.....	21
2.3.3. Nutrition5k.....	21
2.3.4. MalasyanFood11 .....	22
2.3.5. PortugeseMeals .....	22
2.3.6. Fruits 360.....	23
2.3.7. USDA-FNDDS .....	24
<b>3. Realizacija sustava za estimaciju kalorija</b> .....	<b>25</b>
<b>3.1. Struktura i faze implementacije sustava</b> .....	<b>25</b>
<b>3.2. Realizacija procesa standardizacije skupova podataka</b> .....	<b>27</b>
3.2.1. Mape označavaju kategorije.....	28
3.2.2. Priložena je meta datoteka skupa podataka.....	30
3.2.3. Rizici neprovjeravanja sadržaja skupova podataka .....	31
<b>3.3. Ideja cjevovoda</b> .....	<b>31</b>
3.3.1. Učitavač skupova podataka .....	32
3.3.2. Univerzalni One-Hot Encoder.....	34
3.3.3. Optimizacija formata podataka .....	35
3.3.4. TFRecord .....	35
<b>3.4. Duboko učenje i konvolucijske neuronske mreže</b> .....	<b>36</b>
3.4.1. Komponente konvolucijskih neuronskih mreža .....	37
3.4.2. Prijenosno učenje .....	38
3.4.3. Korištenje MobileNetV3 i NASNet modela.....	38
<b>3.5. Pregled modela za estimaciju kalorija</b> .....	<b>39</b>
<b>3.5.1. Korištene biblioteke</b> .....	<b>39</b>

3.5.2. Aktivacijske funkcije, hiperparametri i optimizator .....	40
3.5.3. Pregled arhitekture i slojeva modela .....	41
<b>4. Provedeni eksperiment i dobiveni rezultati .....</b>	<b>44</b>
4.2. Odabrane metode evaluacije .....	45
4.2. Rezultati za pojedinačne skupove podataka .....	46
4.3. Rezultati Kombinacijom dvaju skupova podataka .....	49
4.4. Rezultati Kombinacijom triju skupova podataka .....	51
4.5. Rezultati za cijeli skup podataka .....	53
<b>5. Zaključak .....</b>	<b>57</b>
<b>Literatura.....</b>	<b>58</b>

# Uvod

Svijet je zahvaćen globalnom epidemijom pretilosti. Svjetska zdravstvena organizacija (WHO) izvijestila je da više od 1,5 milijardi odraslih osoba pati od nekog oblika pretilosti. Dokazano najbolji način da se izbjegne prejedanje, koje najčešće dovodi do problema s debljanjem, je praćenje unosa hrane i brojanje kalorija tijekom dana. Glavni problem s ovim pristupom je što mjerenje kalorija unesenih u svakom obroku uvelike ovisi o subjektivnom iskustvu i točnosti osobe koja to radi. Budući da ljudi to najčešće rade sami bez stručne pomoći, moguće je da im se proces zakomplicira što dovodi do greške u procjeni i ponekad preskaču obroke. Nedavno su istraživanja pokazala da korištenje tehnologija poput pametnih telefona može poboljšati liječenje pretilih pacijenata i pacijenata s prekomjernom tjelesnom težinom, pružajući im alternativu brojanju kalorija, što može biti lakše dugoročno rješenje zbog jednostavnosti samog procesa. Ovo rješenje bi omogućilo brojanje kalorija svakoj osobi koja prati dnevni unos kalorija, bilo da se radi o sportskoj karijeri, održavanju trenutne težine ili gubitku male količine viška kilograma. Cilj ovog diplomskog rada je istražiti moć umjetne inteligencije u svrhu rješavanja problema estimacije kalorija, ući dublje u tehnička rješenja koja u pozadini koriste strojno učenje te jedan od tih pristupa testirati tijekom diplomskog rada. Veliki problem većine zadataka strojnog učenja proizlazi iz skupova podataka, te je ovaj rad iznimno fokusiran na proces prikupljanja i standardizacije skupova podataka koji su korisni modelu u estimaciji kalorija. Korišteno je više skupova podataka i raženo je na određenoj količini automatizacije koja bi dovela do lakšeg povećanja novih skupova podataka. Na kraju se prikupljeni i standardizirani skupovi podataka testiraju na algoritmima koji rješavaju problem estimacije kalorija, te se dobiveni rezultati uspoređuju.



# 1. Pregled literature

Tijekom istraživanja korištene su specifične strategije za pretraživanje znanstvenih članaka, koje su navedene u nastavku. Odabran je vremenski period od 2015. do 2024. godine za analizu članaka, s obzirom na napredak u području neuronskih mreža tijekom tog razdoblja. Za pretraživanje znanstvenih članaka navedene su specifične ključne riječi i citatne baze podataka, omogućujući drugima da ponove rezultate. U ovom radu korištene su ključne riječi iz priložene Tablice 1.1.

Tablica 1.1. Ključne riječi korištene prilikom pretraživanja članaka

Ključne riječi	
1.	Calorie estimation AND deep learning
2.	Volume estimation AND deep learning
3.	Image segmentation AND calorie estimation
4.	Object recognition AND calorie estimation

Zatim su korištene sljedeće citatne baze podataka koje osiguraju valjanost podataka te pružaju brzu pretragu i filtraciju članaka.

- Web Of Science
- Scopus

Kako bi se proces pretrage učinio efikasnijim, definirani su kriteriji isključivanja i uključivanja koji će dodatno ubrzati pretragu članaka i eliminirati one koji nisu relevantni. U Tablici 1.2. prikazani su definirani kriteriji.

Tablica 1.2. Kriterij isključivanja i uključivanja korišten prilikom pretraživanja članaka

<b>Kriteriji isključivanja</b>	<b>Kriteriji uključivanja</b>
Nije naveden skup podataka	Priloženi rezultati evaluacije
Nije javno dostupno	Koristi umjetnu inteligenciju
Nisu navedene reference	Napisano je 2015 i kasnije
Nije korišten engleski jezik	Skup podataka je javno dostupan

Pronađena su dva sustavna pregleda iz ovog područja koji su poslužila kao osnova za istraživanje ove teme. Prvi sustavni pregled ovog područja [1] pruža osnovu za cijelu temu procjene kalorija iz različitih perspektiva, a ne samo metoda koje koriste duboke mreže. Drugi sustavni pregled ovog područja [2] napisan je 2024. i pruža dubinsko razumijevanje trenutnog znanja o temi iz perspektive skupova podataka, što je ključno za ovu temu. Oba rada su značajno uštedila vrijeme izrade ovog diplomskog rada jer kvalitetno sažimaju područje istraživanja.

## 1.1. Arhitektura procesa određivanja kalorija

Kako bi se pravovaljano pristupilo temi potrebno je upoznati se s procesom i pojedinim koracima postupka određivanja kalorijskih vrijednosti iz slika. Nijedan pojedinačni rad nije korišten kao osnova za izradu ovog diplomskog rada. Umjesto toga, pojedina znanja su proizašla iz više različitih članaka, što je omogućilo razvoj konceptualne ideje o potrebnim koracima i radnjama za cjelovitu izradu postupka određivanja kalorija iz slika hrane. Npr. u sljedećem opsežnom sistematskom pregledu literature [1], prvi korak odnosi se na izvor podataka, odnosno na način prikupljanja skupova podataka za treniranje. Iako će ovaj aspekt biti implicitno prisutan kroz cijeli rad, glavni fokus ovog diplomskog rada je na korištenju javno dostupnih skupova podataka o hrani koji se mogu pronaći na internetu, kao i na njihovoj primjeni.

U drugom radu [2] mogu se pronaći detaljno analizirani dostupni skupovi podataka hrane na internetu. Proces određivanja kalorijske vrijednosti iz slike hrane može se podijeliti na dva glavna koraka.

- Prepoznavanje objekata
  - Klasifikacija slika i lokalizacija objekata
  - Segmentacija slika
- Procjena volumena/mase
  - Određivanje kalorija

Određivanje što se točno nalazi na slici, osobito kad se radi o više objekata, ključno je područje u strojnome i dubokom učenju poznato kao prepoznavanje objekata. Proces započinje određivanjem položaja objekta te zatim klasifikacijom tog objekta. Kombiniranjem ova dva algoritma postiže se detekcija objekta. Slike dobivene tijekom detekcije objekta još uvijek nisu prikladne za analizu i istraživanje volumena, budući da detekcija objekta identificira okvire popularno zvani *bounding boxes* oko potencijalnih objekata na slici. Da bi se izdvojili željeni dijelovi iz slike ,koriste se metode segmentacije slike.

### **1.1.2. Skupovi podataka za prepoznavanje hrane**

Budući da je prvi korak u radu s većinom problema u strojnom učenju pronalaženje skupova podataka na kojima se temelje sve ostale aktivnosti, identificiran je rad [2] koji učinkovito sistematizira dostupne skupove podataka vezane uz procjenu kalorija i prehranu općenito. Ovaj rad pruža dobar uvid u trenutno dostupna saznanja o toj temi i predstavlja odličnu osnovu za daljnja istraživanja. Stoga se informacije iz tog rada često koriste prilikom tijekom ovog istraživanja, čime se izbjegava potreba za dodatnim naporom u sažimanju već utvrđenih podataka.

### 1.1.3. Kratki pregled trenutno poznatih skupova podataka I njihovog sadržaja

Veliki broj skupova podataka javno je dostupan, što zasigurno omogućava sve veći napredak u području rada sa hranom. Ovi skupovi podataka omogućuju rješavanje problema kao što je klasifikacija hrane, što otvara mogućnosti za brojne aplikacije, uključujući praćenje prehrane, personalizirane prehrambene preporuke i analizu prehrambenih navika, čak i za ideje za recepte hrane.



Slika 1.1. Slike iz skupa podataka UNIMIB2016

U Tablici 1.3. prikazani su neki od često korištenih skupova podataka za prepoznavanje hrane. Svaki skup podataka ima različiti broj slika i klasa te dolazi iz potpuno različitih zemalja. Npr. "UNIMIB 2016" (Slika 1.1) iz 2016. godine sadrži 1027 slika koje se nalaze 73 klase, dok "Yummly-66K" iz 2018. godine sadrži čak 66,615 slika unutar samo 10 klasa.

Tablica 1.3. Usporedba sadržaja više skupova podataka [2]

Skup podataka	Godina	Broj slika/Broj klasa	Država podrijetla
UNIMIB 2016	2016	1027/73	Globalno
ECUSTFD	2017	2978/19	Globalno
Yummly-66K	2018	66615/10	Globalno
Malaysian Food	2018	3300/11	Malezija
KenyanFood13	2019	8174/13	Kenija
Sushi-50	2020	3963/50	Globalno
TKF	2020	1644/34	Indonezija

#### 1.1.4. Izazovi trenutno poznatih skupova podataka

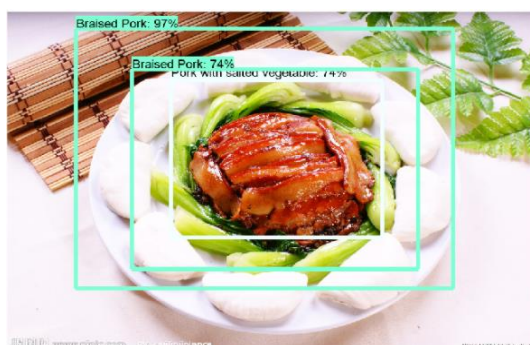
Klasifikacija hrane je problem jednak većini problema u strojnom učenju, s brojnim nizom izazova, posebno kad je riječ o skupovima podataka. Skupovi podataka mogu se značajno razlikovati u broju slika i kategorija, što može utjecati na izbor metoda obrade. Manji skupovi sadrže od 646 do 10.000 slika [1], te se mogu učinkovito obraditi metodama strojnog učenja poput *support vectore machine* or *k-nearest neighbors*. S druge strane, skupovi podataka koji se smatraju velikima, a sadrže desetke tisuća do milijune slika, zahtijevaju napredne algoritme dubokog učenja poput konvolucijskih neuronskih mreža zbog njihove složenosti.

Jedan od glavnih izazova predstavlja odnos između složenosti skupova podataka i preciznosti klasifikacije. Veći skupovi podataka s brojnim kategorijama često otežavaju postizanje visoke točnosti, posebno kada su kategorije hrane nerazmjerno zastupljene. Kada su kategorije dobro uravnotežene i sadržaj velikog skupa podataka odgovara klasifikaciji, rezultati mogu biti izuzetno dobri. Ako neki skupovi podataka sadrže vrlo mali broj slika po kategoriji, time se dodatno otežava postizanje preciznih i generaliziranih rezultata zbog nedovoljnog broja slika za

treniranje u skupu podataka. Geografska distribucija skupova podataka o hrani često nije ravnomjerna, pri čemu mnoge regije i kulinarske tradicije ostaju nedovoljno zastupljene. Ovaj nedostatak raznolikosti može ograničiti primjenjivost modela za klasifikaciju hrane na različitim jelima. Stoga je važno razvijati sveobuhvatnije skupove podataka, rješavati problem neravnomjernosti klasa i pažljivo birati slike koje ulaze u skup podataka.

### 1.1.5. Detekcija objekata pomoću dubokog učenja

Može se koristiti standardna konvolucijska neuronska mreža (engl. CNN – *Convolutional neural network*) za detekciju objekata [3]. Mreža obrađuje slike koje sadržavaju više objekata. Nakon 30.000 iteracija, postignuta je točnost od 97,56%, trenirana na skupu podataka Food101. Na Slici 1.2. je prikazano kako izgleda detekcija pojedenih objekata hrane sa slike.



Slika 1.2. Detekcija objekta nakon 25000 koraka [3]

U Tablici 1.4. uspoređeni su rezultati detekcije objekata za jednostavne komade hrane u usporedbi sa kompleksnim jelima.

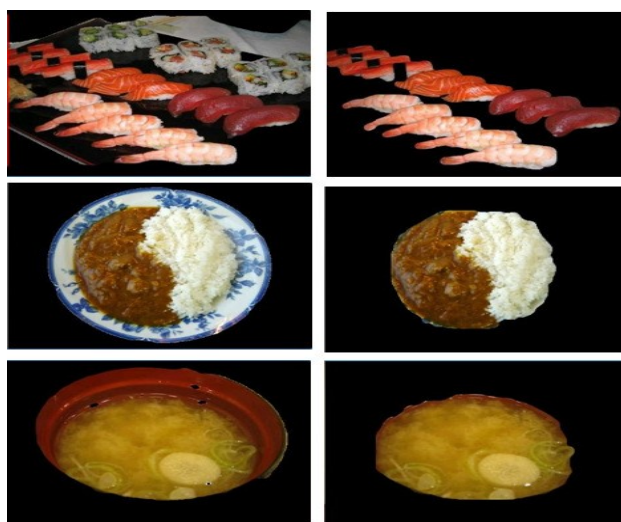
Tablica 1.4. Usporedba CNN-ovih rezultata za otkrivanje objekata pomoću dva različita skupa podataka [3]

Skup podataka	Točnost za jednostavne komade hrane	Točnost za kompleksna jela
UNIMIB 2016	90.80%	93.06%
Customized food plate	85.74%	89.99%

Kao potencijalno poboljšanje u odnosu na konvencionalne CNN-ove, istraživači [4] su koristili FASTER R-CNN za detekciju objekata.

### 1.1.6. Segmentacija slike pomoću dubokog učenja

Nakon identifikacije objekta i postavljanja okvira oko njega, sljedeći važan korak je precizna segmentacija objekta koji se želi obraditi. Postoje dva glavna pristupa za ovaj proces: segmentacija instanci i semantička segmentacija. Ovisno o odabranoj metodi, rezultati se mogu značajno razlikovati. Algoritam za segmentaciju slike obuhvaća kombinaciju više algoritama koji se koriste u različitim fazama. Npr. algoritam MASK R-CNN integrira nekoliko različitih algoritama [5]. ResNet 50 i potpuno konvolucijske mreže (FCN) i dalje se koriste u određenim fazama obrade. Odabiru se namirnice koje imaju fiksnu i poznatu veličinu kao referentne namirnice. Sam proces segmentacije temelji se na mapama istaknutosti s GrabCut algoritmom što se vidi na Slici 1.3. Nadalje, implementacija GrabCut algoritma je opisana u naknadnom radu [7], ističući njegovu primjenu u automatskoj anotaciji segmentacijskih maski.



Slika 1.3. Segmentacijska maska koristeći GrabCut algoritam [3]

U Tablici 1.5. uspoređeni su brojni različiti radovi koji su pružili rezultate za rješavanje problema segmentacije slike.

Tablica 1.5. Usporedba višestrukih rezultata za problem segmentacije slike

Članak	Algoritam	Skup podataka	Tip segmentacije	Točnost
[4]	RefineNet	UNIMIB 2016	semantička	nije navedeno
[5]	Mask R-CNN	COCO	semantička	nije navedeno
[6]	GrabCut	Boiled rice photos	nije navedeno	nije navedeno
[8]	CNN	UNIMIB 2016	semantička	90.80%

### 1.1.7. Metode određivanja volumena hrane

Nakon segmentacije slike, sljedeći korak, određivanje volumena hrane iznimno je bitan u preciznom procesu estimacije kalorija. Razumijevanje svih dostupnih metoda za procjenu kalorija i trenutno znanje o toj temi iskorišteno je iz opsežnog sistematskog pregleda literature [1]. U nastavku su predstavljene neke trenutno istražene metode iz pregleda područja koje se koriste za procjenu volumena hrane:

- Pristup temeljen na stereo-slici je pristup koji podrazumijeva korištenje više kadrova za rekonstrukciju 3D strukture objekata hrane pronalaženjem odgovarajućih piksela između slika i korištenjem vanjskih parametara za reprojekciju piksela iz koordinatnog sustava slike u koordinatni sustav [1].
- Modelni pristup uključuje prethodnu izgradnju oblika uzoraka (matematičkih modela) kako bi se volumen objekata odredio odabirom modela, a zatim skaliranjem i rotacijom modela, što je također poznato kao registracija slike [1].
- Pristup temeljen na dubinskim kamerama koristi dubinske kamere kao što je *Time of flight* (TOF) kamera. Korištenjem pristupa temeljenog na dubinskoj kameri može se dobiti stvarna skala objekta bez potrebe za referentnim objektom kao što je kovanica [1].
- Pristup transformacije perspektive odnosi se na metodu procjene volumena objekta na temelju jedne slike. Korištenjem transformacije perspektive može se dobiti snimka iz ptičje perspektive i napraviti gruba procjena veličine objekta. Ova metoda ne ovisi o



prethodno izgrađenim uzorcima oblika, stoga se obično koristi za procjenu objekata nepravilnog oblika [1].

- Pristup temeljen na dubokom učenju koristi isključivo algoritme dubokog učenja u procjeni volumena. Nekoliko istraživačkih radova predložilo je korištenje jedne RGB slike za stvaranje dubinske mape. Voksel reprezentacija je korištena za prikaz dubinske mape, a volumen se može procijeniti brojanjem broja zauzetih vokselâ. Međutim duboko učenje nudi i niz drugih opcija za rješavanje problema estimacije volumena [1].

Kada se uspoređuju različite metode za procjenu volumena hrane, stereo metoda se ističe svojom visokom točnošću (oko 90%) jer omogućuje korištenje više slika za rekonstrukciju 3D strukture hrane. Ova metoda je teška za svakodnevnu primjenu zbog potrebe za specijaliziranim tehnikama, što je nepraktično i zahtijeva dodatno vrijeme. Pristup temeljen na modelu pokazuje učinkovitost od 85% za standardne oblike hrane, poput modela jabuke, ali je manje pouzdan za nepravilne oblike zbog ograničenog broja unaprijed definiranih modela. Metode koje koriste dubinsku kameru pružaju točne rezultate (oko 92%) bez potrebe za referentnim objektima, no njihova široka primjena je ograničena zbog visoke cijene i nedostatka dostupne opreme. Metode perspektivne transformacije mogu fleksibilno procijeniti nepravilne oblike (oko 80%), ali su općenito manje precizne u usporedbi s metodama koje koriste više slika ili senzora. Metoda dubokog učenja pokazala se najperspektivnijom (95%) zbog svoje sposobnosti prilagodbe hrani različitih oblika i tekstura s visokom točnošću. [1]

Iako postoji potreba za obradom velikih količina podataka, metode dubokog učenja će se nastaviti istraživati zbog njihove prilagodljivosti i mogućnosti rada bez dodatne opreme.

### **1.1.8. Određivanje kalorija**

Nakon što se utvrdi volumen hrane, bez obzira na metodu koja je korištena za njegovu određivanje, konačno izračunavanje kalorija zahtijeva masu hrane u gramima. Ta se masa dobiva iz formule gustoće i volumena hrane. Gustoća je unaprijed poznata i može se pronaći u tablici gustoće za razne namirnice. Jednadžba za izračunavanje mase (1):

$$\text{ukupna masa} = \text{volumen} \times \text{gustoća} \quad (1)$$

Kada se dobije masa hrane, izračunavanje njezinog kalorijskog sadržaja može se obaviti jednostavnom formulom. Ukupna količina kalorija izračunava se množenjem kalorija, za 100 grama s utvrđenom masom određenog prehrambenog proizvoda prema sljedećoj jednadžbi:

$$\text{ukupne kalorije} = (100 \text{ kalorija}) \times \text{ukupna masa} \quad (2)$$

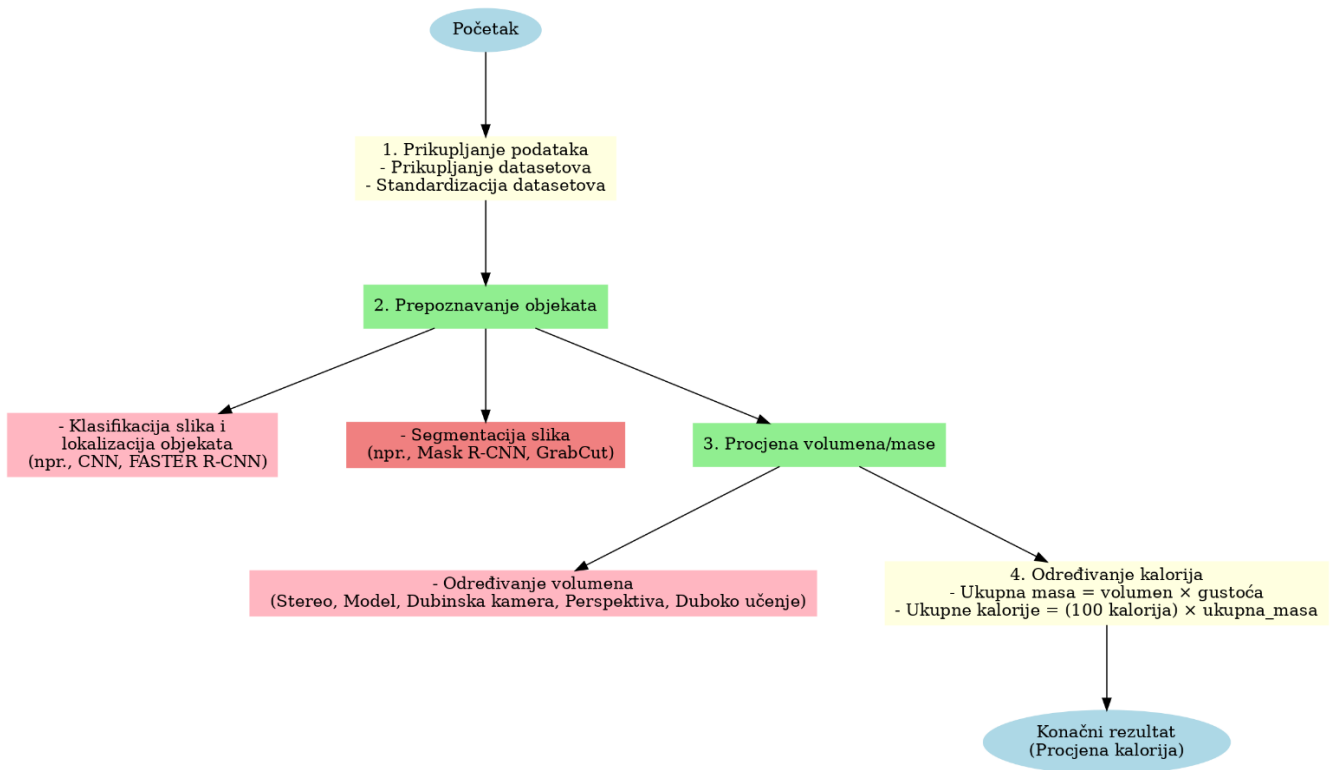
U Tablici 1.6. je konačni rezultat procjene kalorija kao završni rezultat cijelog procesa procjene kalorija.

Tablica 1.6. Usporedba višestrukih rezultata za problem procjene kalorija

Rad	Skup podataka	Točnost
Jaswanthi et al. [18]	UNIMIB 2016	95.21%
Popy et al. [11]	UNIMIB 2016	93.06%
Chen et al. [21]	Chinese Food	96.76%
Jian et al. [15]	FooDD	87.02%
Frank et al. [22]	Food dataset	97.6%

### 1.1.9. Dijagram toka

Prikazani dijagram toka nije kopija logike jednog istraživačkog rada, već je zapravo sinteza različitih koraka navedenih u više relevantnih istraživačkih radova. Cilj je bio povezati korake i metodologije iz različitih izvora kako bi se stvorila sveobuhvatna ideja te zatim i sama konstrukcija strukture dijagrama toka određivanja kalorija iz slika hrane. Članci iz kojih je dobivena ideja za stvaranje ovog dijagrama uključuju radove poput [1, 5, 6, 7, 8] i [11]. Iz ovih radovima, postignuto je bolje razumijevanje glavnih komponenti cijelog procesa estimacije kalorija te to obuhvaća pojedine korake kao što su prepoznavanja objekata [3], tehnika segmentacije slika [5] i konačna tehnika procjene volumena hrane [1]. Na Slici 1.4. prikazan je dijagram toka koji je poslužio u ovom radu za izradu cjelokupnog sustava za estimaciju kalorija.



Slika 1.4. Dijagram toka za problem estimacije kalorija

## 1.2. Korišteni Algoritmi i dosadašnji rezultati na specifičnim skupovima podataka

Kako bi se kasnije moglo uspoređivati rezultate, u nastavku su navedeni različiti skupovi podataka te algoritmi koji su primijenjeni na njima, kako bi se vidjelo kako isti algoritam može dobiti slične ili drugačije rezultate na različitim skupovima podataka.

Algoritam DenseNet-201 je primijenjen na dva različita skupa podataka: Food101 i UEC Food-256. Na skupu podataka Food101, algoritam je postigao točnost od 80.63%, dok je na skupu UEC Food-256 postigao nešto bolju točnost od 81.12%. Ovo pokazuje da iako je struktura algoritma identična, sam skup podataka može imati značajan utjecaj na performanse modela. Razlike u veličini skupa podataka, varijaciji slika, te broju klasa značajno doprinose ovim razlikama u točnosti. Daljnje analize također pokazuju kako različiti algoritmi djeluju na istom skupu podataka. Skup Food101 je testiran koristeći više različitih arhitektura, to su Inception-ResNet-V2 (74.11%), DCNN (77.4%), i EfficientNet-B4 (91.50%). EfficientNet-B4 pokazuje

vrlo dobru točnost u odnosu na druge algoritme, što ukazuje na to da modernije arhitekture značajno mogu poboljšati performanse klasifikacije slika. U usporedbi s Food101 skupom, na skupu UEC Food-256, rezultati pokazuju općenito nižu točnost. GoogLeNet postiže tek 54.5% na UEC Food-256, dok na Food101 postiže 53.96%. Iako su rezultati slični, DCNN na UEC Food-256 postiže točnost od 67.6%, što je primjetno niže od 77.4% na Food101, što može sugerirati da UEC Food-256 sadrži više varijabilnosti ili da je složeniji za klasifikaciju nego Food101.

Tablica 1.7. Usporedba više algoritama na dva specifična skupa podataka

<b>Referenca</b>	<b>Skup podataka</b>	<b>Algoritam</b>	<b>Točnost</b>
[12]	Food101	Inception-ResNet-V2	74.11%
[14]	Food101	DCNN	77.4%
[15]	Food101	WISeR	90.27%
[14]	Food101	GoogLeNet	53.96%
[17]	Food101	EfficientNet-B4	91.50%
[13]	Food101	ResNet-50	80.84%
[13]	Food101	DenseNet-201	80.63%
[13]	UEC Food-256	DenseNet-201	81.12%
[16]	UEC Food-256	DCNN	67.6%
[14]	UEC Food-256	GoogLeNet	54.5%
[13]	UEC Food-256	Inception-ResNet-V2	81.54%
[15]	UEC Food-256	WISeR	83.15%

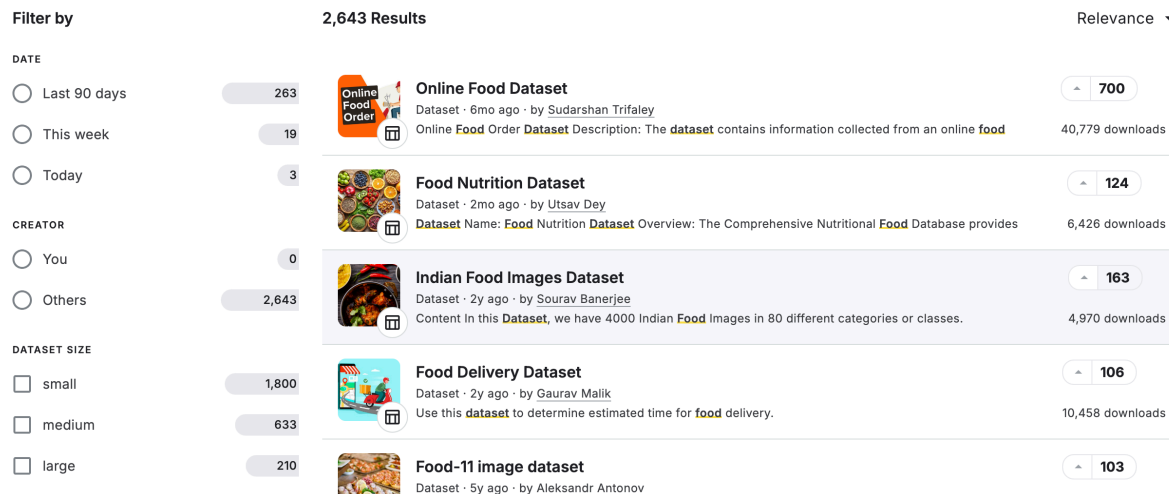
## 2. Pronalaženje, istraživanje i standardizacija skupova podataka za hranu

Proces izrade kreće određivanjem skupova podataka na kojima će konačni model estimacije biti istreniran. Dostupni skupovi podataka koji se spominju kroz istražene članke su kvalitetan izvor za izgradnju procesa, međutim, istraženi su i dodatni izvori podataka te načine prikupljanja podataka. Postoji više opcija kako se podaci mogu prikupljati te stvoriti skup podataka. Prva opcija je "prikupljanje podataka samostalno" koja može biti vrlo vremenski zahtjevna i ovaj tip prikupljanja podataka se najčešće koristi u velikim tvrtkama. Unatoč velikoj vremenskoj zahtjevnosti samostalno prikupljanje podataka često može biti najbolje rješenje jer omogućava potpunu kontrolu nad procesom i dobivanje najrelevantnijeg skupa podataka kojim se može poboljšati učinkovitost modela strojnog učenja budući da će taj oblik prikupljanja podataka ukloniti nepotrebne podatke i zadržati smislene podatke potrebne za razvoj kvalitetnog modela. Druga opcija je pronaći dostupne skupove podataka na Internetu. Postoje dvije vrste podataka na Internetu – skupovi podataka otvorenog pristupa i komercijalni skupovi podataka. Pronalaženja skupova podataka na Internetu je mnogo brže, no uz određena ograničenja kao što je mogućnost da podaci mogu biti irelevantni za specifične potrebe ili da neće postojati potrebna fleksibilnost u formatu podataka. Postoji i nekoliko drugih opcija, kao što su *web scrapping* te korištenje podataka iz više neslužbenih internetskih izvora. Iako mogu biti vrlo učinkovite, ove metode prikupljanja podataka često zahtijevaju dodatne obrade i moguće je naići i na pravne probleme. U ovom radu korišteni su skupovi podataka otvorenog pristupa dostupni na Internetu, uz neke dodatne postupke i obrade poput standardizacije skupova podataka. Standardizacija je važna jer omogućava modelu da trenira na podacima koje može učinkovito koristiti prilikom treniranja na određenim algoritmima. Konkretno je prikazano kako razlike među skupovima podataka mogu utjecati na učinkovitost procesa. Osim toga demonstrirano je koliko je izazovno postići standardizaciju kada se koriste javni skupovi podataka sa Interneta i kako bi uvođenje određenih standardizacijskih normi moglo učiniti proces korištenja dostupnih skupova podataka s Interneta lakšim.

## 2.1. Izvori skupova podataka na Internetu

Na Internetu postoji velika količina izvora preko kojih se mogu pretraživati skupovi podataka. Neki izvori nude javno dostupne skupove podataka, a neki nude samo komercijalne skupove podataka. Cilj ovog paragrafa je prikazati koji su izvori korišteni prilikom istraživanja dostupnih podataka o hrani na Internetu. Istaknute su platforme s kojih su preuzeti skupovi podataka koji su korišteni u ovom diplomskom radu.

Jedna od najvećih i najpoznatijih baza podataka za strojno učenje je Kaggle. Na ovoj platformi može se pronaći velika količina resursa za strojno učenje koji sadrže kod, ali i velika količina raznih skupova podataka. Tijekom pretrage podataka o hrani pronađeno je 2643 rezultata. Iako svi rezultati nisu relevantni za temu ovog diplomskog rada, prikazana je značajna količina dostupnih resursa na platformi. Također treba istaknuti jednu od glavnih prednosti Kaggle-a, a to je mogućnost korištenja dodatnih filtera koji omogućuju sužavanje pretrage i uklanjanje skupova podataka koji nisu od interesa. Kaggle kao što se vidi i na Slici 2.1. nudi rangiranje skupova podataka prema raznim kriterijima, poput ocjena korisnika, broja preuzimanja i drugih relevantnih kriterija.



The screenshot shows the Kaggle dataset search results page. On the left, there are filter options under 'Filter by'. The main area displays a list of datasets with their respective icons, titles, descriptions, and download counts. The datasets listed are:

Dataset Name	Downloads
Online Food Dataset	700
Food Nutrition Dataset	124
Indian Food Images Dataset	163
Food Delivery Dataset	106
Food-11 image dataset	103

Slika 2.1. Prikaz dostupnih značajki web stranice Kaggle

Osim Kagglea, pregledani su i ostali izvori skupova podataka, uključujući platforme kao što su DrivenData, DataHub, OpenML i drug3. Svaka navedena platforma pruža velik broj skupova podataka i ostalih resursa koji su od koristi svima koji se bave strojnim učenjem. Međutim, kao

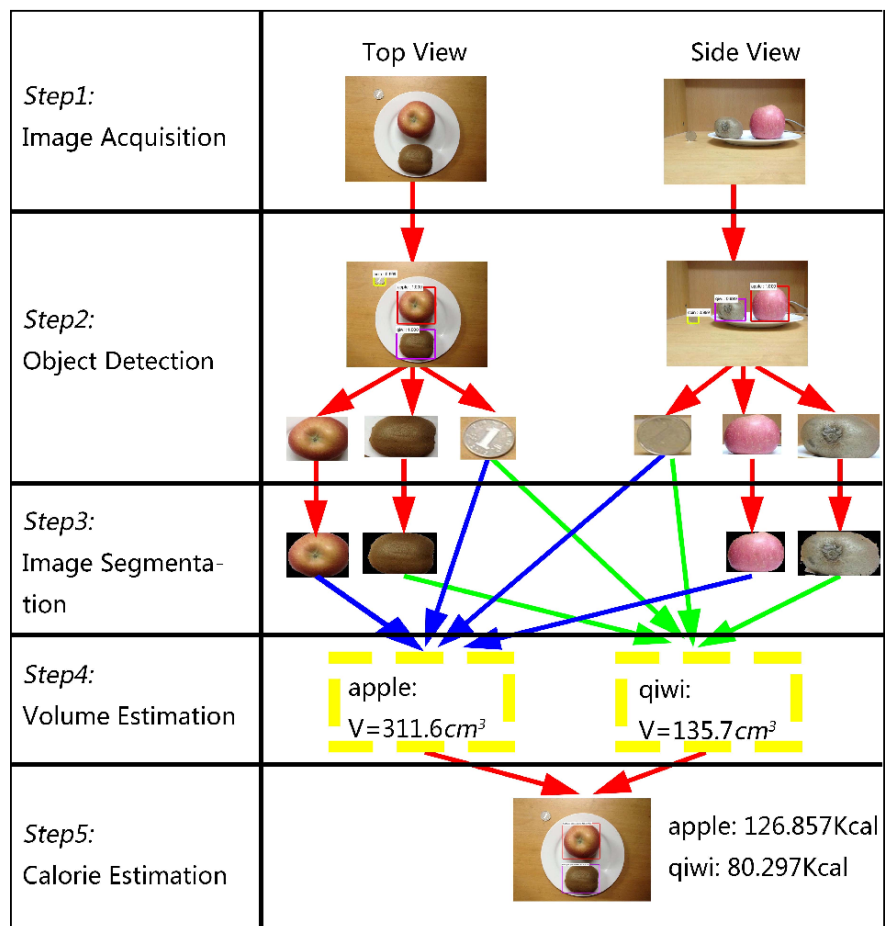
najpopularniji izvor, Kaggle ima najsuperiornije funkcionalnosti i najveću bazu korisnika. Svi resursi su javno dostupni, tako da je Kaggle korišten kao primarni izvor podataka.

Iz članka [2] i istraživanja raznih GitHub repozitorija s ciljem dubljeg razumijevanja procesa izrade ovog diplomskog rada, pronađeni su dodatni izvori skupova podataka. Npr. u spomenutom članku [2] opisuje se web stranica čiji je cilj sistematizirati sve postojeće skupove podataka na jednom mjestu prema državama iz kojih dolaze. Svaki skup podataka na toj stranici popraćen je kratkim opisom i poveznicom koja vodi do izvornog izvora, gdje se skup može preuzeti. Tijekom istraživanja te stranice identificirano je nekoliko dodatnih neslužbenih izvora podataka.

## **2.2. Dostupni skupovi podataka na Internetu i njihove osnovne značajke**

Nakon dugotrajne pretraga raznih skupova podataka koji su navedeni u znanstvenim člancima, ali i skupova podataka sa drugih izvora podataka opisanih u prijašnjem paragrafu, zaključeno je da se sadržaji skupova podataka međusobno razlikuju u određenoj mjeri da se to može odraziti na smjer standardizacije skupova podataka. Počevši od najintuitivnijeg kriterija, same veličine skupa podataka, postavlja se pitanje koliko veće količine podataka zapravo povećavaju vrijednost skupa podataka ili potencijalno smanjuju smisao skupa podataka. Već od samog početka moguće je svrstati skup podataka u određene skupine. Primjerice, manji skup podataka može biti strogo precizan i pružati puno informacija o određenim uzorcima te je koristan za specifične primjene. Što se tiče velikih skupova podataka, obično je korisno za rješavanja probleme klasifikacije ako je skup podataka iznimno velik jer sadrži veliku količinu slika iste vrste hrane. Svaki pojedinačni skup podataka na neki je način specifičan, može se koristiti u drugoj primjeni i treba ga tako tretirati. Proučavanjem skupova podataka zaključeno je da se većina njih uglavnom koristi za rješavanje problema klasifikacije. Npr. skup podataka Food101 pruža veliki broj potpuno različitih slika pite od jabuka i koristi se samo za identifikaciju da se radi o piti od jabuka, što je klasičan problem klasifikacije. Nasuprot tome, skup podataka ECUSTFD [18] pruža veliki broj slika jabuka iz različitih kutova s kalibracijskim objektima,

što je vidljivo na Slici 2.2, što je korisnije za precizne studije na temu određivanja volumena koje se može postići u laboratoriju.



Slika 2.2. Korištenje jednog komada jabuke i referentnog objekta [18]

Pošto pristup koji ima većina skupova podataka uključuje veliku količinu slika hrane, potrebno je predložiti neke standarde korištenja takvih skupova podataka u procesu estimacije kalorija. Nameću se dva pristupa, sve slike standardizirati na referentnu vrijednost od 100 grama, što bi često moglo biti iznimno krivo pretpostavljati da svako jelo ima 100 grama ili puno bolja pretpostavka pretpostaviti da jelo na slici sadrži jednu porciju te istražiti dodatni napredak tog pristupa. U nastavku se nalazi Tablica 2.1. koja je bila korisna prilikom istraživanja skupova podataka te donošenja određenih zaključaka o tipovima skupova podataka s kojima se radi u ovom diplomskom radu.



Tablica 2.1. Sistematizacija karakteristika analiziranih skupova podataka

Skup podataka	Jedna slika za 1 jelo	Više slika istog jela različite veličine jela	Meta data sastava	Sadrži kalorije
UEC-FOOD256	ne	da	Imena mapa su kategorije	ne
Nutrition5K	Više slika dobiveno rastavljanjem videa na frame-ove	ne	Nutritivne vrijednosti Imena mapa su kategorije	da
Recipes5k	ne	da	Samo kategorije	ne
Malaysia Food 11	da	ne	Samo kategorije	ne
ECUSTFD resized	Više za 1	da	Nutritivne vrijednosti Kategorije masa	ne
SlowDeepFood	Više za 1	da	Imena mapa su kategorije	ne
UEC FOOD 256	Više za 1			
UPMC FOOD101			Kategorije	
ETHZ FOOD101	Više za 1	da	sastojci	ne
Chinese FoodNet	Više za 1	da	Kategorije	ne
Food524DB	Više za 1	da	Kategorije	ne

### 2.2.1. Pristup estimacije jedne porcije jela na slici

Kod pristupa estimacije jedne porcije najčešće se radi s velikom količinom skupova podataka različite veličine, geografskih obilježja te kvalitete slika. Takvi skupovi podataka su najdostupniji na Internetu, a uvođenjem raznih metoda standardizacije moguće ih je iskoristiti za različite primjene, u ovom slučaju za primjenu estimacije kalorija. Za većinu takvih skupova podataka primijećeno je da su jedini dostupni podatci kategorije hrane koje su korisne prilikom klasifikacije. Podaci poput mase hrane u gramima i volumena pojedinog jela nisu dostupni. Nisu poznati ni podaci o kalorijama te nutritivnim vrijednostima. Kako bi se iskoristio potencijal opisanih skupova podataka, moguće je metodama standardizacije povezati pojedino jelo s dostupnim tablicama standardnih porcija na internetu i pomoću toga napraviti model za estimaciju kalorija. Model bi mogao imati vrlo lošu točnost, no mogao bi biti dovoljno dobar pokušaj iskorištavanja dostupnih informacija za rješavanje problema estimacije kalorija. Pretpostavka je da su veličine porcija približno slične, te ako je jelo dobro klasificirano, trebali bi se postići približno točni podaci o nutritivnim vrijednostima pojedinog jela. Kao unaprjeđenje modela, mogu se prepoznavati pojedini sastojci jela kako bi se jela što bolje klasificirala te kako bi greška bila što manja. Moguće je iskoristiti i tehnike poput detekcije objekta kako bi se pribrojale različite porcije jela npr. ako je više jabuka, model treba prepoznati količinu jedinki jabuka. Ako je na slici dupla porcija mesa, model klasificira da postoji više mesa na slici. Taj pristup bi se olakšalo postojanje varijanti kvalitetno klasificiranih skupova podataka koji razlučuju različite porcije pojedinog jela. Druga opcija je razviti dovoljno napredne algoritme koji kvalitetno prepoznaju pojedine sastojke jela te prema tome izračunavaju konačni iznos nutritivnih vrijednosti. Preporuka za izradu skupova podataka koji bi pomogli u estimaciji kalorija je sljedeća: razdvojiti porcije pojedinog jela na mape tako da svaka mapa sadrži što sličnije slike hrane sa sličnijim porcijama iz raznih kutova. Npr. ako neko jelo postoji s više komada mesa ili s manje, bilo bi korisno da su razdvojeni u različite mape, te ako je moguće, veće i manje porcije potpuno razdvojiti u različite mape te napraviti kvalitetnu meta datoteku koja jasno daje do znanja što ta mapa predstavlja. Potrebno je u iste mape ne stavljati totalno različite slike istog jela jer to nikako ne pomaže modelu, nego ga zbunjiva i značajno snižava točnost estimacije kalorija. Na Slici 2.3. prikazana je razlika porcije pite od jabuka i cijele pite od jabuka.



Slika 2.3. Primjer razlike između porcije pite od jabuka i jela cijele pite od jabuka

## 2.3. Karakteristike i sadržaj odabranih skupova podataka

Nakon analize velike količine skupova podataka iz različitih izvora, odabrano je pet skupova podataka za rad sa slikama hrane, kao i dva skupa podataka koji se fokusiraju na nutritivne vrijednosti. Pri odabiru skupova podataka nije postavljen strogi kriterij koji bi zahtijevao identične karakteristike svih podataka. Svaki skup podataka je ručno provjeren kako bi se osiguralo da je prikladan za korištenje, a konačni odabir bio je djelomično nasumičan. U nastavku su predstavljeni odabrani skupovi podataka.

### 2.3.1. FOOD 101

U članku [9] predstavljen je skup podataka FOOD101 koji obuhvaća 101 kategoriju hrane, pri čemu svaka kategorija sadrži 750 slika za treniranje i 250 slika za testiranje. Autori su ručno obradili testni skup podataka, dok skup podataka za treniranje sadrži određenu razinu "šuma" među kategorijama, što znači da su slike s različitim vrstama hrane označene jednom klasom. Ručno očišćeni testni podaci, koji se sastoje od 25.000 slika, nazivaju se Food-101-CUR i koriste se kao referentni skup za proširenje skupa podataka. Za validaciju se koristi 10% neuređenih podataka za obuku, dok se preostalih 90% (označenih kao Food-101-UNCUR) koristi za augmentaciju podataka prilikom treniranja dubokih modela. Na Slici 2.4. prikazane su slike iz skupa podataka food101.



Slika 2.4. Slike iz skupa podataka food101

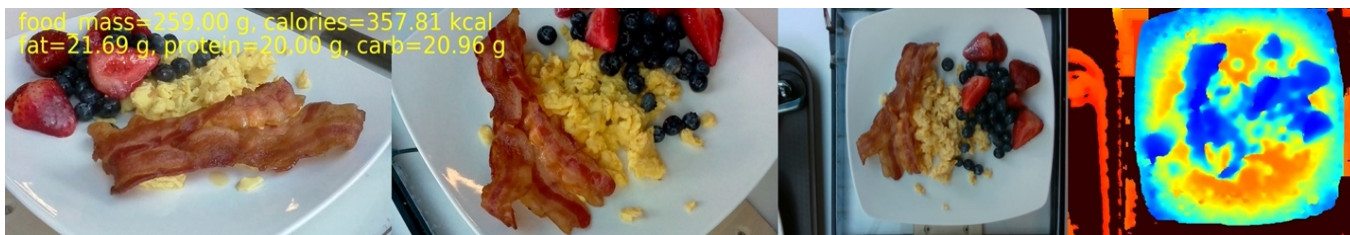
### 2.3.2. Recipes5k

U članku [10] predstavljen je skup podataka Recipes5k, razvijen za prepoznavanje sastojaka, koji sadrži 4.826 jedinstvenih recepata. Svaki recept uključuje sliku i popis sastojaka, a ukupno se u skupu nalazi 3.213 različitih sastojaka, s prosječnim brojem od 10 sastojaka po receptu. Svaki recept nudi alternativni način pripreme jednog od 101 tipa hrane iz skupa podataka Food101, što omogućava prepoznavanje varijacija sastojaka unutar istih jela. Kako bi se osigurala kvalitetnija distribucija, oko 50 alternativnih recepata za svaku od 101 klase ravnomjerno je raspoređeno u skupove za treniranje, validaciju i testiranje. Jedan od izazova u radu s 3.213 sirovih sastojaka je prisutnost mnogih potklasa (npr. 'narezana rajčica' ili 'umak od rajčice'), koje predstavljaju specifičnije verzije šireg pojma (npr. 'rajčica'). Kako bi se pojednostavio ovaj problem, predložena je verzija skupa podataka koja izostavlja previše deskriptivna obilježja (poput 'narezana' ili 'umak'), čime je broj sastojaka smanjen na 1.013, a ovi pojednostavljeni sastojci su korišteni za daljnje evaluacije.

### 2.3.3. Nutrition5k

Nutrition5k [11] je skup podataka razvijen s ciljem unapređenja razumijevanja nutritivnog sastava hrane putem analize vizualnih podataka. Ovaj napredni skup podataka obuhvaća otprilike 5.006 stvarnih jela zabilježenih u Google-ovim kafićima uz pomoć specijaliziranog sustava za skeniranje. Svako jelo prati video materijal snimljen iz različitih kutova, RGB-D slike snimljene odozgo kada su dostupne, kao i detaljan popis sastojaka koji uključuje masu pojedinih komponenti, ukupnu masu jela i precizno označene nutritivne vrijednosti. Ovaj skup podataka omogućuje obuku algoritama koji mogu točno procijeniti kalorijsku vrijednost i sadržaj makro nutrijenata (masti, proteina i ugljikohidrata) u složenim jelima. Predložena je osnovna metoda

za poboljšanje točnosti nutritivnih procjena korištenjem dostupnih podataka o dubini slike. Nutrition5k dolazi s odvojenim skupovima za treniranje i testiranje te uključuje skripte za evaluaciju regresije nutritivnih vrijednosti. Ovaj skup podataka, zajedno s povezanim istraživanjima, objavljen je s ciljem poticanja daljnjih inovacija u razumijevanju prehrane kroz vizualne podatke. Na Slici 2.5. prikazane su slike iz skupa podataka Nutrition5k.



Slika 2.5. Slike iz skupa podataka Nutrition5k

#### 2.3.4. MalasyanFood11

Skup podataka MalaysianFood11, poznat i kao myFood11, predstavlja kolekciju slika hrane fokusiranu na malezijsku kuhinju. Ovaj skup podataka obuhvaća 11 različitih kategorija jela, a slike su organizirane u odgovarajuće mape. Među kategorijama nalaze se popularna malezijska jela poput Nasi Goreng, Ayam Goreng i Laksa, kao i mnoge druga. Skup podataka je podijeljen na 80% za obuku i 20% za testiranje, a slike su označene u PASCAL VOC formatu, koji se često koristi za zadatke prepoznavanja objekata.

#### 2.3.5. PortugeseMeals

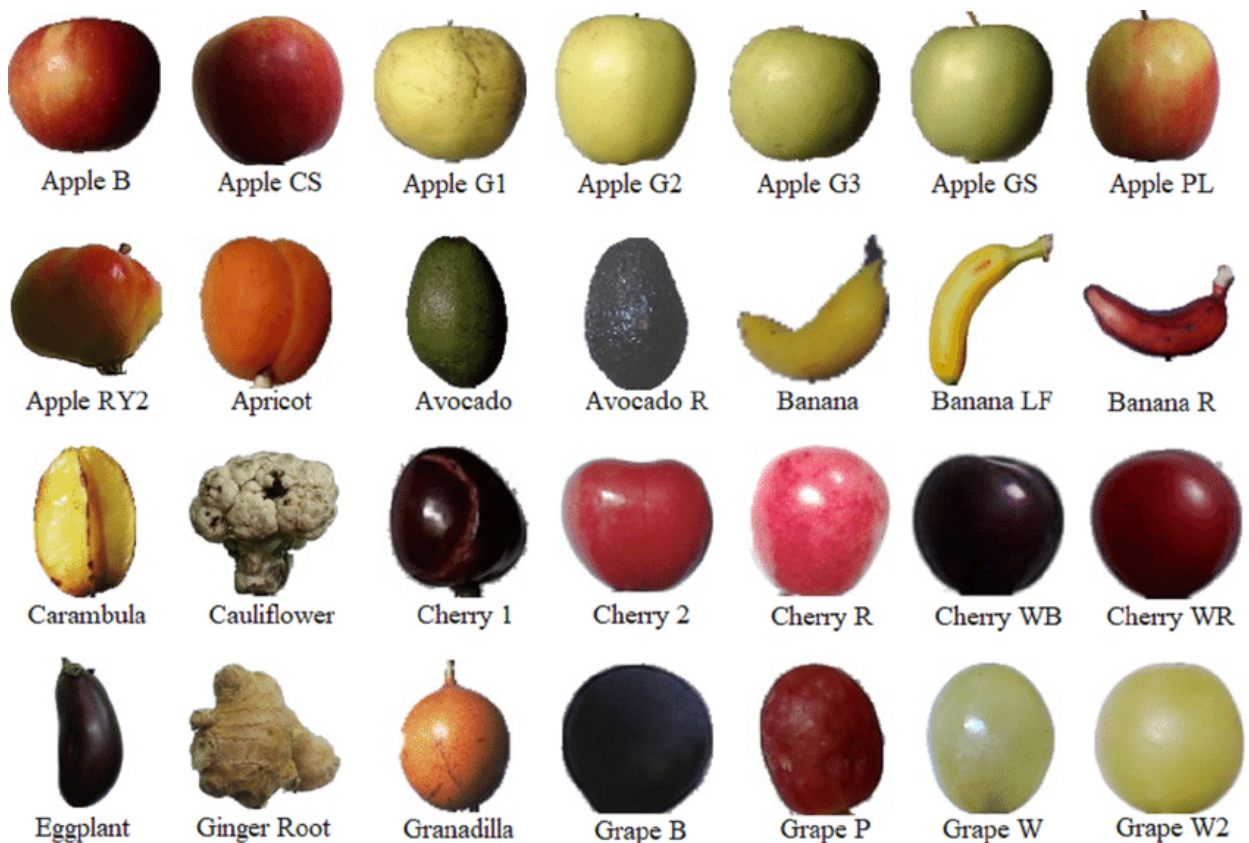
Ovaj skup podataka [39] kreiran je za razvoj sistema prepoznavanja portugalskih jela i naknadnih nutritivnih izračuna, posebno za pomoć dijabetičarima u odabiru hrane. Skup sadrži 6771 sliku raspoređenu u 23 kategorije, s posebnim fokusom na portugalska jela, koja su dijelom prikupljena iz Food-101 skupa podataka, dok su ostala jela ručno prikupljena sa Google Slika i aplikacija poput Zomato i TripAdvisor. Cilj je da ovaj skup postane najveći dostupan Internet resurs za portugalska jela. Na Slici 2.6. prikazane su slike iz skupa podataka Portuguese Meals.



Slika 2.6. Slike iz skupa podataka Portuguese Meals [40]

### 2.3.6. Fruits 360

Fruits360 [42] je jedan od najpoznatijih i najčešće korištenih skupova podataka za klasifikaciju voća u području strojnog učenje. Ovaj skup podataka [37] obuhvaća slike raznih vrsta voća i povrća snimljenih u kontroliranim uvjetima, što osigurava dosljednost prikaza objekata. Ukupno je dostupna 141 klasa, pri čemu svaka klasa predstavlja određenu vrstu voća ili povrća. Ove klase uključuju širok spektar voća, od jabuka i banana do egzotičnijih vrsta poput karambole i rambutana. Skup podataka sadrži više od 90.000 slika, koje su podijeljene u skupove za treniranje i testiranje, gdje je oko 70.000 slika za treniranje omogućujući precizno uvježbavanje modela za klasifikaciju slika. Sve slike su standardizirane na rezoluciju od 100x100 piksela, što osigurava ujednačenost ulaznih podataka i smanjuje varijabilnost koja bi mogla utjecati na performanse modela. Zbog svoje veličine, raznolikosti i kvalitete, Fruits360 se često koristi u istraživačkim projektima, diplomskim radovima i industriji za razvoj i testiranje modela klasifikacije slika. Na Slici 2.7. prikazane su slike iz skupa podataka Fruits360.



Slika 2.7. Slike iz skupa podataka Fruits360 [38]

### 2.3.7. USDA-FNDDS

Baza podataka o hrani i hranjivim tvarima za studije o prehrani (FNDDS) [43] predstavlja ključni izvor informacija o hranjivim vrijednostima hrane i pića koji su zabilježeni u okviru Nacionalnog istraživanja o zdravlju i prehrani (NHANES), poznatom kao „Što jedemo u Americi“ (WWEIA). [35] Njena osnovna svrha je pretvaranje podataka o konzumiranim namirnicama i napicima iz WWEIA i NHANES u grame te određivanje njihove hranjive vrijednosti. Budući da FNDDS generira datoteke o unosu hranjivih tvari za WWEIA i NHANES, više nije potrebna odvojena procjena unosa hranjivih tvari iz anketa. Baza podataka se redovito ažurira svake dvije godine, u skladu s objavama podataka o prehrani iz WWEIA i NHANES, pri čemu je svaka verzija prilagođena specifičnim istraživačkim razdobljima, osiguravajući točan prikaz hrane, pića i hranjivih vrijednosti relevantnih za to vrijeme.

### 3. Realizacija sustava za estimaciju kalorija

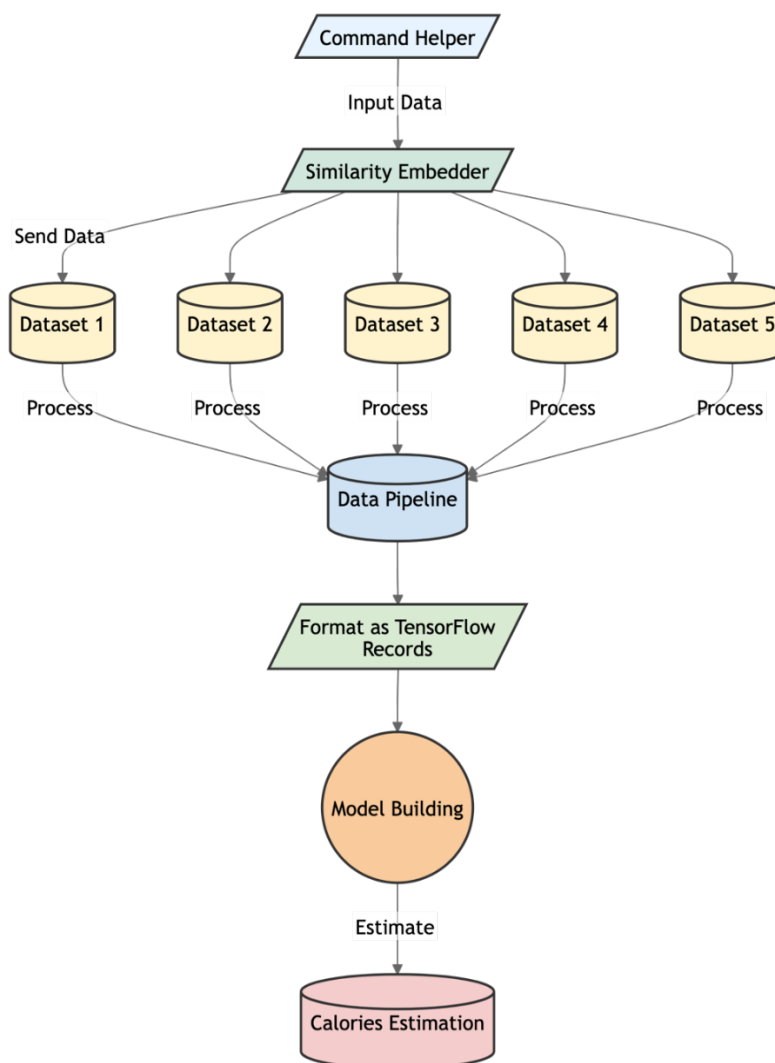
Nakon detaljne analize literature i znanstvenih članaka, stječe se konceptualna ideja o potrebnim koracima za realizaciju procesa estimacije kalorija. Pronađeni su različiti prijedlozi za primjenu algoritama u tim koracima, no kao što je to često slučaj, znanstveni članci ne sadrže detaljne tehničke opise izrade pojedinih koraka potrebnih za realizaciju cjelokupnog projekta. Kako bi se identificirali konkretni koraci u pisanju koda, korišteni su izvori poput koda generiranog putem ChatGPT-a, znanje o Pythonu te ideje iz drugih prethodnih izrađenih projekata. Sustav je u potpunosti razvijen unutar Google Colab okruženja, koristeći programski jezik Python, kao i dostupne biblioteke i značajke Google Colab-a. Cjelokupni kod je organiziran u obliku Colab bilježnica. S obzirom na to da se u ovom diplomskom radu posebna pažnja posvećuje standardizaciji podataka, važno je istaknuti ideje i koncepte koji su primijenjeni kako bi se postigla kvalitetna standardizacija odabranih skupova podataka. Drugi dio tehničkog segmenta fokusira se na implementaciju procesa procjene kalorija korištenjem algoritama dubokog učenja.

#### 3.1. Struktura i faze implementacije sustava

Struktura sustava je osmišljena s jasno definiranim rasporedom mapa i datoteka, s ciljem uspostave konačnog cjevovoda koji omogućava automatsko izvođenje cijelog procesa u organiziranom okruženju. U sustavu postoje specifični dijelovi programskog koda koji osiguravaju interoperabilnost između datoteka i Google Colab bilježnica. Bilježnica "CommandHelper" prikazana na Slici 3.1. koristi se za izvršavanje komandi koje se pokreću jednom ili za obavljanje raznih procesa i akcija nad datotekama i zip arhivama. Najčešće se koristi za povezivanje s platformom Kaggle i za preuzimanje i raspakiranje skupova podataka. Ova bilježnica je stoga inter operabilna s ostalim bilježnicama. Bilježnica "*Similarity embedder*" namijenjena je izradi tablice koja obuhvaća raznolike kategorije hrane, sastojke te sve važne nutritivne vrijednosti i informacije o kalorijama. U ovom diplomskom radu korištene su meta datoteke iz skupova podataka USDA-FNDDS i nutrition5K u procesu izgradnje tablica. Na Slici 3.1. svaki skup podataka prikazan je odvojeno, što se može primijetiti i u strukturi sustava, jer su organizirani u posebne mape, pri čemu svaki skup podataka ima svoju bilježnicu za pred procesiranje. Bilježnica "data pipeline" predstavlja korak koji prethodi izgradnji modela, u



kojem se svi podaci generirani iz različitih skupova podataka prikupljaju i spremaju u format koji je prihvatljiv za model. U ovom procesu postoje i dodatni koraci koji su ključni za izgradnju automatizacije i velikog cjevovoda. Završna faza, u kojoj se modeli implementiraju i izvršavaju, naziva se "*build model*". U ovoj fazi se procjenjuju svi elementi pripremljeni u prethodnim koracima, uključuju se arhitekture algoritama i njihovo treniranje, te pozivaju se razne funkcije i koraci automatizacije razvijeni unutar cjevovoda.



Slika 3.1. Dijagram tehničkog sustava

## 3.2. Realizacija procesa standardizacije skupova podataka

Kao što je prikazano na Slici 3.1. u prethodnom poglavlju, svaki skup podataka ima svoju mapu i pripadajuću Google Colab bilježnicu koja sadrži proces standardizacije. Ovaj pristup je neophodan jer svaki skup podataka na početku zahtijeva individualno razmatranje kako bi se uklonile potencijalne razlike i problemi. Iako su razvijene funkcije koje omogućuju automatizaciju ovog koraka, potpuno automatiziranje procesa ostaje izuzetno izazovno. Cilj je postići potpunu automatizaciju u budućnosti, no postoje brojne nesavršenosti u različitim segmentima koje otežavaju ostvarenje tog cilja. U vezi s ovim pitanjem identificirano je mnogo potencijalnih problema, što ukazuje na to da potpuna automatizacija možda nije ni moguća. Tokom izrade Google Colab bilježnica za standardizaciju, istaknuta su dva osnovna tipa skupova podataka koja se mogu obraditi.

- Mape označavaju kategorije
- Priložena je meta datoteka skupa podataka

Budući da se većina skupova podataka može svrstati u ova dva tipa, neovisno koje podatke sadrže, od samog početka standardizacije razmatrane su funkcije koje obrađuju jedan od ta dva tipa skupova podataka i uspješno s njima rade, naravno uz manja prilagođavanja koja su potrebna za svaki pojedini skup podataka. Time se usmjerava ka većoj automatizaciji, olakšanom dodavanju novih skupova podataka i bržem testiranju tih skupova. To podrazumijeva ubrzan proces integracije novih podataka, od samog početka pa sve do kraja cjevovoda i modela obrade. Prije početka s procesom automatizacije, važno je definirati što automatizacija i standardizacija svakog skupa podataka trebaju postići. U nastavku su sažeto predstavljene očekivani rezultati koje bi Google Colab bilježnice trebale generirati za svaki pojedini skup podataka. Svaki očišćeni skup podataka trebao bi sadržavati sljedeće parametre:

- **ID/File Name:** Jedinstveni identifikator ili naziv datoteke za svaki skup podataka. Ovaj parametar omogućava pronaći odgovarajuće slike za taj skup podataka.
- **Category:** U kontekstu ovog rada, kategorije su jela.
- **Calorie (kcal):** Sadržaj kalorija izražen u kilokalorijama.

- Carbohydrate (g): Sadržaj ugljikohidrata u gramima.
- Protein (g): Sadržaj proteina u gramima.
- Fat (g): Sadržaj masti u gramima.
- Ingredients: Popis sastojaka koji se nalaze u jelu.

Nakon čišćenja i obrade podataka, potrebno je za svaki skup podataka izraditi meta datoteku u CSV formatu. Ova meta datoteka mora obavezno sadržavati sve prethodno navedene parametre. U kontekstu ovog diplomskog rada, važno je napomenuti da se sve datoteke koje su uspješno obrađene premještaju u zajedničku mapu pod nazivom "*full dataset*". Međutim, korištenjem zasebnih meta datoteka moguće je izolirati pojedinačne skupove podataka, što je korisno kada je potrebno istražiti njihov utjecaj na konačni rezultat.

### 3.2.1. Mape označavaju kategorije

U nastavku je opisan proces obrade možda jednostavnijeg, ali vrlo često viđenog problema u praksi, situacije kad skup podataka koristi mape kao kategorije. Prvi obrađen skup podataka je FOOD101, na kojem je postavljen osnovni kostur koji se može primijeniti na bilo koji drugi sličan tip skupova podataka. U ovom slučaju se ne primjenjuju *embedder*-i, već se uspoređuju kategorije skupa podataka koji ne sadrži nutritivne vrijednosti i kalorije s kategorijama skupa podataka za koje te informacije dostupne. Tijekom obrade skupa podataka s potpunim nutritivnim vrijednostima, na kraju je odlučeno korištenje *recipes5k*. Proces započinje identifikacijom kategorija iz skupa podataka *food101* koje su 80% ili više slične kategorijama iz skupa podataka *recipes5k*. Nakon čega se te kategorije mapiraju, a u konačnoj meta tablici se koriste slike iz *food101* skupa podataka, dok se kategorije i ostale informacije koriste iz *recipes5k* tablice. Isti model je primijenjen na još dva skupa podataka: *Malaysianfood1* i *Portuguesemeals*.

Problem s ovim pristupom odmah se očituje kada meta podaci koji se analiziraju ne sadrže traženu kategoriju ili nijednu sličnu, što je slučaj u ovom diplomskom radu zbog iznimno ograničenih skupova podataka. U takvim okolnostima gotovo je nemoguće odrediti nepoznatu kategoriju, pa se mora odlučiti između mapiranja na nešto slično što odgovara situaciji, ili proširivanja tablice vlastitim podacima, što može dovesti do unosa netočnih informacija u

proces. Još jedan zanimljiv problem u istraživanju domicilnih skupova podataka, poput Malaysianfood1 i Portuguesemeals, je to što sadrže jela koja su nepoznata svjetskim organizacijama. Njihovo mapiranje je nepouzđano, a dodavanje tih podataka u tablicu nije dobra praksa ako se želi postići pouzđan model. Dok su jela poput hamburgera, pizze i sličnih često lako obradiva, ona značajno doprinose konačnom skupu podataka za klasifikaciju. U nastavku na Slici 3.2. je prikazano kako se nepoznata jela mapiraju s određenim poznatim vrijednostima.

```
category_mapping = {  
    "fried_noodles": "gyoza",  
    "kaya_toast": "croque_madame",  
    "laksa": "hot_and_sour_soup",  
    "mixed_rice": "paella"  
}
```

Slika 3.2. Mapirane kategorije za skup podataka MalasyanFood11

Kod malezijske kuhinje može se pronaći neka jela koja su neviđena u drugim svjetskim kuhinjama. U portugalskom skupu podataka postoje jela koja odgovaraju nekim svjetskim specijalitetima, ali zbog specifičnosti prijevoda potrebno ih je dodatno mapirati.

```
category_mapping = {  
    "aletria": "bread_pudding",  
    "arroz_cabidela": "chicken_curry",  
    "bacalhau_bras": "fish_and_chips",  
    "bacalhau_natas": "creme_brulee",  
    "batatas_fritas": "french_fries",  
    "bolo_chocolate": "chocolate_cake",  
    "cachorro": "hot_dog",  
    "caldo_verde": "french_onion_soup",  
    "cozido_portuguesa": "beef_tartare",  
    "croissant": "croque_madame",  
    "esparquete_bolonhesa": "spaghetti_bolognese",  
    "feijoada": "bean_tartare",  
    "francesinha": "club_sandwich",  
    "gelado": "chocolate_ice_cream",  
    "hamburguer": "hamburger",  
    "jardineira": "ratatouille",  
    "nata": "creme_brulee",  
    "ovo": "omelette",  
    "pasteis_bacalhau": "crab_cakes",  
    "tripas_moda_porto": "beef_tripe",  
}
```

Slika 3.3. Mapirane kategorije za skup podataka PortugeseFood

Važno je napomenuti da je korišten skup podataka pod nazivom FRUITS-360, koji sadrži različite vrste voća, a pripada kategoriji "mape kao kategorije". Budući da voće nije kompletan obrok, potrebno je izvršiti modifikacije kako bi se odredila kalorijska vrijednost pojedinačnih sastojaka, Npr. jabuka se može smatrati sastojkom. Ovo je samo jedan od mnogih različitih scenarija koji se mogu pojaviti u praksi u okviru ovako složenog modela rada sa hranom, gdje postoji velika raznolikost nepredvidivih situacija.

### 3.2.2. Priložena je meta datoteka skupa podataka

U ovom slučaju, priložene su slike hrane zajedno s meta datotekom koja povezuje te slike s određenim vrijednostima. Konkretno, u slučaju Recipes5k, dostupna su mapiranja slika kao oznake za obuku i testiranje. Budući da se radi o receptima, dostupne su kategorije različitih namirnica i liste sastojaka.. No za izradu potpune tablice nedostaju podatci o kalorijama i nutritivnim vrijednostima. Kako bi se izradila sveobuhvatna tablica koju bi svaki skup podataka trebao generirati na kraju procesa standardizacije, potrebno je razviti *embedder*-e koji će obraditi tekstualne datoteke i usporediti ih, koristeći kosinusnu sličnost sa Slike 3.4. za identifikaciju sličnih pojmova. U ovom slučaju uspoređuju se slične kategorije i jela iz tablica recipes5k, USDAFNDDS i nutrition5K. Kada se identificira najbliže jelo, određene informacije iz tih tablica se spajaju kako bi se generirala glavna meta datoteka. Priprema se ista tablica kao recipes5k, koja se na kraju koristi za pomoć drugim skupovima podataka u pronalaženju sličnih jela. Ograničenja ovog procesa uključuju i količinu jela koja su dostupna. Razvijanje ovog automatiziranog pristupa može biti korisno za bilo koji drugi skup podataka koji ima sličnu strukturu.

```
for ingredient in ingredients_list:
    embedding = Ingredients.embed([ingredient])
    (
        fndds_similarity,
        fndds_index,
        fndds_category,
    ) = Ingredients.get_most_similar_from_fndds(embedding)
    (
        nutrition5k_similarity,
        nutrition5k_index,
    ) = Ingredients.get_most_similar_from_nutrition5k(embedding)
    row = []
    if fndds_similarity > 0.7 and fndds_similarity > nutrition5k_similarity:
        ingredient_nutrition = Ingredients.get_ingredient_nutrient_from_fndds(
            fndds_category, fndds_index
        )
```

Slika 3.4. Prikaz dijela koda u kojem se koriste funkcije istraživanja sličnosti

### 3.2.3. Rizici neprovjeravanja sadržaja skupova podataka

Veliki problem ovog procesa je što nije moguće pregledati sadržaj svakog skupa podataka i osobno se uvjeriti da svaka slika ima smisao i zadovoljava potrebnu kvalitetu za rad. Zbog toga u mapama postoji velika količina slika koje možda nisu od velike koristi za procjenu kalorija, ali bi mogle biti korisnije za sam proces klasifikacije jela. Primjer je sljedeća slika hamburgera koja se često poslužuje uz pomfrit. Na ovoj slici može se intuitivno zaključiti da pomfrit pripada hamburgeru, zbog čega se razvijaju složeniji modeli koji moraju prepoznati određene sastojke i označiti da jelo nije samo hamburger, već da uključuje i pomfrit. U ovom ograničenom broju kategorija, to jelo će se uvijek svrstati kao hamburger, iako uključuje i pomfrit.



Slika 3.5. Slika hamburgera i pomfrita iz skupa podataka PortugueseFood

### 3.3. Ideja cjevovoda

Cjevovod (eng. *Data Pipeline*) [19] predstavlja skup automatiziranih procesa koji omogućuju unos, validaciju, obradu i integraciju različitih skupova podataka prije njihove upotrebe u modelima. Korištenje cjevovoda osigurava dosljednost, efikasnost i preciznost u radu s raznolikim podacima. Kako se navodi u radu [19]; "u situacijama gdje se koristi više heterogenih skupova podataka, implementacija robusnog cjevovoda nije samo korisna, već i neophodna za osiguranje kvalitete podataka i dosljednosti u analitičkim rezultatima".

U slučaju modela za procjenu kalorija, potrebno je obraditi slike hrane, tekstualne opise sastojaka i kategorije, kao i numeričke vrijednosti nutrijenata. Svaka od ovih vrsta podataka

zahtijeva poseban pristup obrade, što može otežati proces integracije. Korištenjem cjevovoda, moguće je automatizirati standardizaciju formata podataka za model kroz integraciju različitih izvora informacija. Time se ne smanjuje samo vrijeme potrebno za pripremu podataka, već se i osigurava da su svi podaci u konzistentnom formatu prije nego što uđu u proces treniranja modela. Razina automatizacije koja se ovdje može postići je ograničena, te nije moguće stvoriti cjelovit cjevovod koji bi od nestrukturiranih podataka generirao savršeno strukturiran ulaz za model. Nakon završetka s odvojenim koracima standardizacije, za svaki skup podataka koristi se posebna Google Colab bilježnica koja priprema podatke u uniformnom formatu za daljnju obradu. Primjeri sadržaja u cjevovodu mogu uključivati izradu funkcija koje učitavaju datoteke i provode potrebne transformacije kako bi se dobio ulazni format za model. Svaka vremenska ušteda i automatizacija nekog od koraka procesa je od iznimne važnosti u velikim sustavima gdje je potrebno brzo testiranje. U kontekstu ovog diplomskog rada, prilikom rada s više skupova podataka cilj je što brže obraditi podatke te ih poslati modelu na testiranje u odgovarajućem formatu. Svaki korak automatizacije, posebno ukupni cjevovod, postaje izuzetno važan i značajno ubrzava proces.

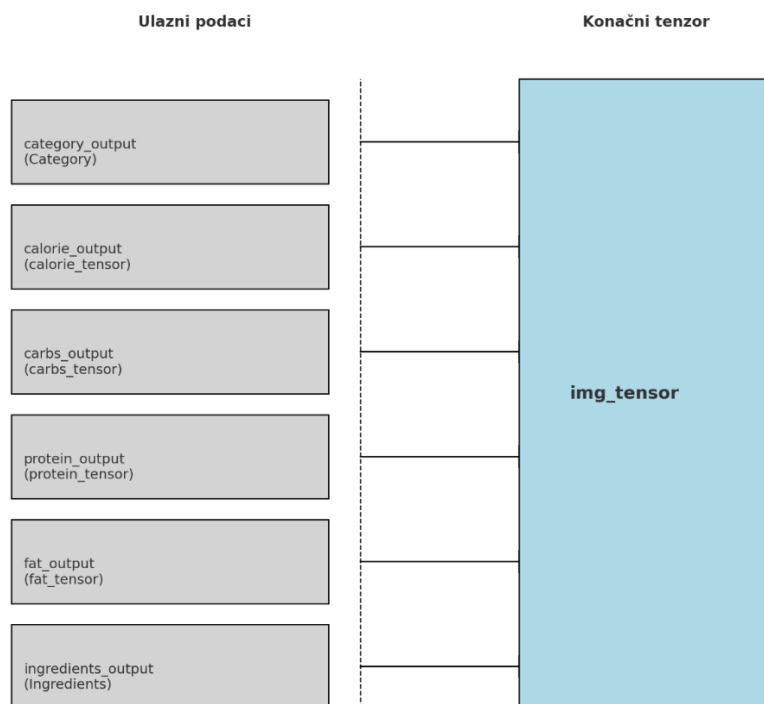
U nastavku će biti prikazane ključne komponente realizirane implementacijom cjevovoda. Umjesto mjerenja brzine i postignutog ubrzanja, naglasak je stavljen na korisničko iskustvo i na to kako integracija sljedećih komponenti čini proces implementacije novih rješenja bržim i efikasnijim. Ključne komponente ostvarene u cjevovodu su:

- Učitavač skupova podataka
- Univerzalni onehotencoder
- Optimizator formata podataka – tf record

### **3.3.1. Učitavač skupova podataka**

Učitavač podataka započinje proces učitavanjem meta podataka iz odgovarajuće datoteke. Meta podaci za svaki skup podataka pohranjeni su odvojeno. Ovaj proces uključuje i standardizaciju dimenzija slika, što je jedna od učestalih praksi u radu sa slikama u strojnom učenju. Učitavač podataka omogućava i ekstrakciju svih jedinstvenih kategorija hrane i sastojaka iz skupa podataka, kao i pozicija slika. Ova funkcionalnost je ključna za razlikovanje pojedinih stupaca skupa podataka, jer će se na kraju koristiti kao različite komponente u problemu procjene

kalorija. Jedna od glavnih funkcija učitavača je izvođenje operacija obrade slika hrane za korake treniranja modela. Važan korak u tom procesu je učitavanje slika i njihovo pretvaranje u nizove. Standardizirani formati slika omogućuju kompatibilnost s različitim arhitekturama neuronskih mreža, čime se povećava fleksibilnost i primjenjivost modela. U kontekstu procjene kalorija, važno je utvrditi što svaki redak treba sadržavati nakon obrade i pretvorbe u potrebni format za model, kao i kako ga pravilno pohraniti. Kao optimalno rješenje nameće se sljedeće. Tenzori [21] su višenamjenski matematički objekti koji generaliziraju skalarne vrijednosti (nulte dimenzije), vektore (prve dimenzije) i matrice (druge dimenzije) na višedimenzionalne prostore. U strojnom učenju, posebno u dubokom učenju, tenzori se koriste za predstavljanje i manipulaciju podacima u različitim dimenzijama, kao što su slike, zvučni zapisi ili tekst. Tenzori omogućavaju jednostavnu manipulaciju podacima kroz optimizirane matematičke operacije, što omogućava efikasno učenje i izvođenje modela. Budući da je potrebno kodirati različite podatke kao što su kategorije, sastojci i nutritivne vrijednosti, a već na početku postoje raznolike tekstualne i numeričke informacije koje je potrebno povezati sa slikama, tenzori se nameću kao jedino adekvatno rješenje. Na Slici 3.6 prikazano je šta svaki kodirani podatak treba obuhvatiti, a to predstavlja jedan tenzor koji se smatra jednim podatkom u procesu treniranja modela.

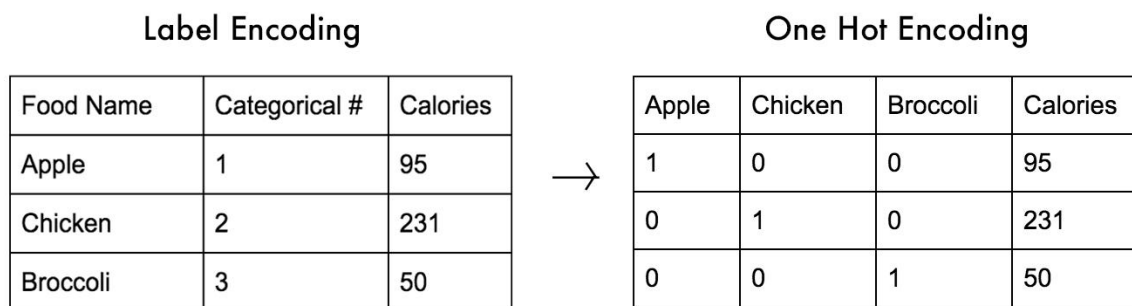


Slika 3.6. Prikaz kreiranja jedinstvenog tenzora za jednu sliku i njegov sadržaj



### 3.3.2. Univerzalni One-Hot Encoder

*One-hot Encoding* je važna tehnika u strojnome učenju koja omogućuje pretvaranje kategorijskih podataka u numerički format, što je ključno za rad s većinom algoritama strojnog učenja. Algoritmi poput neuronskih mreža, regresije i stabala odlučivanja zahtijevaju numeričke ulaze za svoje operacije, stoga je neophodno pretvoriti tekstualne ili kategorijske podatke u numeričke vektore. [22]. *One-hot Encoding* svaku kategoriju predstavlja kao vektor s binarnim vrijednostima, pri čemu je svaka kategorija mapirana na točno jednu poziciju u vektoru koja ima vrijednost jedan, dok su sve ostale pozicije nula. Ovakav pristup omogućuje algoritmima da razlikuju različite kategorije korištenjem numeričkih zapisa [23]. U ovom diplomskom radu, *One-hot Encoder* igra ključnu ulogu u efikasnoj obradi raznih kategorija hrane i sastojaka iz nekoliko različitih skupova podataka. S obzirom na to da svaki skup podataka sadrži specifične kategorije i sastojke, korisno je imati univerzalni *One-Hot Encoder* koji može objediniti sve te kategorije. Primjenom univerzalnog *One-Hot Encodera* [24] postiže se dosljedna reprezentacija podataka, što smanjuje mogućnost grešaka prilikom treniranja modela na različitim skupovima podataka. Implementacija univerzalnog *One-Hot Encodera* od posebne je važnosti u okviru automatizacije i razvoja sustava za obradu podataka. Univerzalni *One-Hot Encoder* omogućuje centralizirano i standardizirano kodiranje, čime se smanjuje potreba za prilagodbama ili ručnim intervencijama prilikom dodavanja novih skupova podataka ili kategorija [25]. U ovom sustavu koji koristi cjevovod, univerzalni *One-Hot Encoder* igra ključnu ulogu u omogućavanju skalabilnosti. Novi skupovi podataka mogu se lako integrirati u postojeći sustav bez potrebe za prepravkom dijelova koda vezanih uz kodiranje kategorija, što značajno pojednostavljuje održavanje sustava i omogućuje brzu implementaciju novih modela ili unapređenje postojećih, a da pritom ne dođe do kompromitacije dosljednosti ili kvalitete podataka.



Slika 3.7. Primjer onehotenkodiranja sa hranom i kalorijama [1]

### 3.3.3. Optimizacija formata podataka

Rad s velikim skupovima podataka, posebno u području hrane, donosi brojne izazove u obradi i analizi informacija. Kada se radi s tisućama slika različitih prehrambenih proizvoda, postoje ograničenja u maksimalnoj veličini datoteka, a složenost struktura podataka se povećava. Ako podaci nisu pravilno formatirani ili pripremljeni ili se ne uklapaju prilikom pohrane u neodgovarajuće vrste dokumenata, mogu se javiti problemi poput produženog vremena treniranja modela, neučinkovite upotrebe računalnih resursa i poteškoća u interpretaciji rezultata. Npr. jedan od problema koji se može pojaviti je da je model obučen za prepoznavanje slika prehrambenih proizvoda i njihovo povezivanje s podacima o hranjivim vrijednostima. Ova faza modela vjerojatno će trajati duže od očekivanog, posebno kada su podaci pohranjeni u formatima poput CSV-a ili JSON-a, čak i kada su uspostavljene odgovarajuće strukture za Excel tablice. Ipak, takve metode su obično prikladne za čišćenje i učitavanje podataka, ali se ne preporučuju za velike zadatke obrade podataka u strojnom učenju, jer produžuju vrijeme učitavanja i obrade što može značajno ometati napredak treniranja modela. Postoji i potreba za formatom podataka koji omogućava značajno smanjenje dimenzija podataka, kao i njihovu upotrebu na različitim uređajima uz pomoć unaprijed definiranih klasa učitavača.

### 3.3.4. TFRecord

Kako bi se riješili izazovi iz prošlog poglavlja, uveden je format TFRecord, optimizirani format za pohranu i dijeljenje velikih skupova podataka. Sadrži datoteku u binarnom formatu koja je dizajnirana za pohranjivanje i rukovanje raznim količinama podataka kao što su slike, tekst ili brojevi. Brzina operacija čitanja i pisanja podataka glavna je prednost tog formata koja uvelike smanjuje vrijeme potrebno za obuku modela strojnog učenja. Brzina čitanja i pisanja podataka predstavlja ključnu prednost ovog formata, značajno skraćujući vrijeme potrebno za obuku modela strojnog učenja.

U TFRecord formatu podaci se također koriste i pohranjuju u serijaliziranom obliku, najčešće uz pomoć YAML serijalizacijskih nizova što olakšava njihov kasniji rad s podacima. Istraživanja [15] pokazuju da format TFRecord može unaprijediti izvedbu modela strojnog učenja, posebno kada se radi o obradi velikih količina podataka. Ova faza procesa obično uključuje pretvaranje mnogih vrsta podataka (slika, oznaka, podataka o hranjivim

vrijednostima) u binarne nizove koji se naknadno pohranjuju u datoteke TFRecord. Rad s podacima postaje lagan i ubrzava se izračune tijekom obuke modela. U ovom diplomskom radu, koji se bavi problematikom procjene kalorija te automatizacijom tog procesa, optimizacija formata podataka je vrlo bitna za uspješnu implementaciju modela strojnog učenja. Radi se s različitim skupovima podataka koji uključuju slike hrane i njihove nutritivne vrijednosti. Kako bi se osigurala konzistentnost i efikasnost u obradi podataka, pohranjeni su skupovi podataka korištenjem TFRecord formata. Time je omogućen brži pristup podacima i smanjeno vrijeme obrade, što je od presudne važnosti za trening modela koji uključuju veliki broj slika i nutritivne informacije. Korištenje TFRecord formata omogućuje paralelno učitavanje podataka, što je smanjilo vrijeme potrebno za pripremu i obradu. Korišteni podaci se jednostavno mogu prenijeti na bilo koji drugi uređaj, što omogućava testiranje modela na različitim uređajima koristeći dostupne informacije spremljene u binarnim datotekama.

### **3.4. Duboko učenje i konvolucijske neuronske mreže**

Jedna od ključnih tehnologija korištenih u ovom radu je umjetna inteligencija. Tijekom proteklog desetljeća duboko učenje postalo je vodeća tehnologija u području umjetne inteligencije, posebno u obradi slike. Duboko učenje, kao podskup strojnog učenja, koristi višeslojne neuronske mreže za modeliranje složenih obrazaca u podacima. Ova metoda omogućava računalnim sustavima da prepoznaju složene obrasce u podacima bez potrebe za ručnim definiranjem karakteristika, što predstavlja značajan napredak u odnosu na klasične pristupe strojnog učenja [26].

Jedan od najuspješnijih tipova dubokih neuronskih mreža za obradu slike su konvolucijske neuronske mreže (engl. *Convolutional Neural Networks*). CNN-ovi su se pokazali izuzetno efikasnim u prepoznavanju obrazaca na slikama, zahvaljujući svojoj arhitekturi koja je inspirirana ljudskim vizualnim sustavom. CNN-ovi automatski uče hijerarhijske karakteristike iz ulaznih slika, počevši od osnovnih elemenata kao što su rubovi i tekture u površinskim slojevima, a zatim se postupno usmjeravaju na složenije oblike i objekte u dubljim slojevima. Ovaj pristup obradi omogućava visoku preciznost u zadacima kao što su prepoznavanje objekata, klasifikacija slika i analiza složenih vizualnih podataka, što je ključno u suvremenim primjenama strojnog učenja. [27].

### 3.4.1. Komponente konvolucijskih neuronskih mreža

Konvolucijske neuronske mreže predstavljaju ključnu arhitekturu u dubokom učenju jer omogućuju preciznu analizu i prepoznavanje značajki unutar slika. CNN-ovi su inspirirani načinom na koji ljudski vizualni sustav obrađuje informacije, te su dizajnirani da automatski uče značajke iz ulaznih podataka kroz niz slojeva koji transformiraju piksele u visoko apstraktne reprezentacije [26]. CNN arhitektura sastoji se od nekoliko ključnih komponenti koje surađuju kako bi omogućile učinkovit proces učenja i generalizacije iz vizualnih podataka.

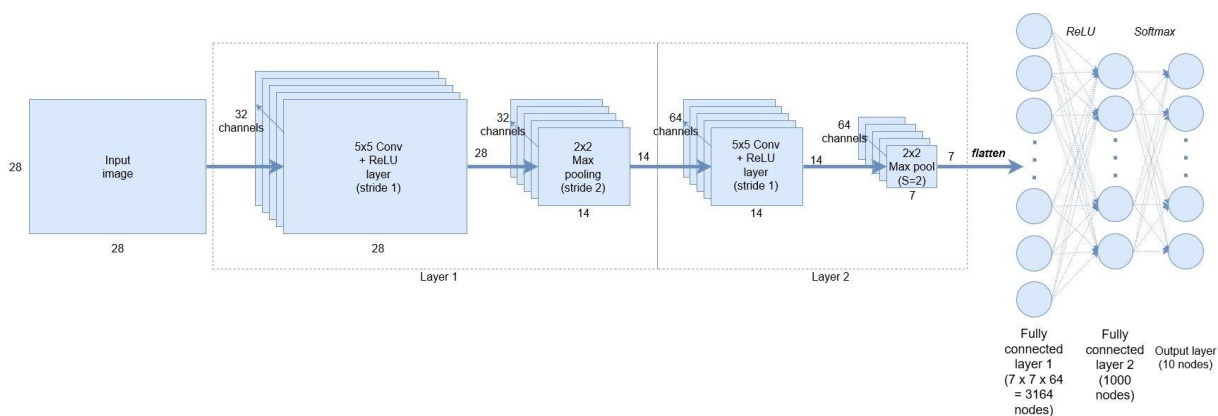
Prva komponenta CNN-a su konvolucijski slojevi, koji čine temelj modela. U tim slojevima koriste se filteri, unaprijed definirane matrice poznate kao kerneli, koji se primjenjuju na ulazne podatke kako bi se izdvojili lokalni obrasci poput rubova, tekstura i osnovnih oblika iz slike. Ovi filtri generiraju značajke koje se još zovu *feature maps* te predstavljaju različite aspekte vizualnih informacija i omogućuju dodatne transformacije.

Korištenjem tehnika kao što su koraci (*stride*), koji određuju pomak kernela, i popunjavanje (*padding*), koje radi na očuvanju rubne informacije slike, može se kontrolirati smanjenje prostorne dimenzije slike, a istovremeno poboljšati učinkovitost učenja modela. Važno je istaknuti da duboke CNN arhitekture mogu sadržavati brojne konvolucijske slojeve, pri čemu se složeniji obrasci prepoznaju u dubljim slojevima mreže [27].

Nakon konvolucijskih slojeva, *pooling* slojevi igraju ključnu ulogu u smanjenju dimenzionalnosti značajki dobivenih konvolucijom. Najčešće korištena metoda je *max pooling*, koja zadržava maksimalnu vrijednost unutar svakog podprostora slike. Osim *max pooling*-a, koristi se i *average pooling*, koji izračunava prosječnu vrijednost unutar podprostora. *Pooling* slojevi ne samo da olakšavaju smanjenje složenosti modela, već također omogućuju apstrakciju informacija, što olakšava kasniju klasifikaciju slika [55].

Potpuno povezani slojevi predstavljaju završnu fazu u arhitekturi konvolucijskih neuronskih mreža (CNN), gdje se značajke iz prethodnih slojeva koriste za donošenje konačnih odluka. U ovim slojevima, svaki neuron je povezan sa svim neuronima iz prethodnog sloja, što omogućava mreži da kombinira naučene značajke i stvara precizne predikcije. Ovaj korak omogućava modelu integraciju informacije iz konvolucijskih slojeva i formiranje zaključaka o cjelokupnoj slici. Posljednji sloj obično koristi *softmax* aktivacijsku funkciju, koja pretvara rezultate u

vjerojatnosti za različite kategorije, čime se omogućava klasifikacija objekata unutar slika [35]. Osim ovih osnovnih komponenti, CNN-ovi često uključuju dodatne tehnike poput normalizacije i *dropout*-a kako bi se poboljšala generalizacija modela i spriječilo prekomjerno učenje popularno zvano *overfitting*. Normalizacija omogućuje ujednačavanje distribucije podataka unutar svakog sloja, dok *dropout* nasumično isključuje određeni postotak neurona tijekom treniranja, čime se smanjuje rizik od prekomjernog učenja modela [56]. U nastavku na Slici 3.8. prikazana je jedna arhitektura neuronske mreže sa svim opisanim slojevima u prethodnim paragrafima.



Slika 3.8. Primjer cjelokupne neuronske mreže

### 3.4.2. Prijenosno učenje

Prijenosno učenje (engl. *transfer learning*) predstavlja jednu od najmoćnijih tehnika u dubokom učenju, posebno u situacijama kada su podaci ograničeni. Temeljna ideja ovog pristupa je korištenje modela koji je prethodno treniran na velikom i raznolikom skupu podataka, kao što je ImageNet, te njegovo fino podešavanje za specifičnu zadaću na manjem skupu podataka. Ova metoda omogućava iskorištavanje već usvojenih značajki, čime se smanjuje potreba za velikim količinama podataka i skraćuje vrijeme potrebno za treniranje [29].

### 3.4.3. Korištenje MobileNetV3 i NASNet modela

U ovom radu, prijenosno učenje primijenjeno je pomoću unaprijed treniranih modela MobileNetV3 i NASNet. Ovi modeli, koji su prvotno trenirani na milijunima slika iz ImageNet

skupa podataka, koriste se kao temelj za izgradnju sustava za klasifikaciju hrane i procjenu nutritivnih vrijednosti.

MobileNetV3 je model dizajniran za mobilne aplikacije, optimiziran za postizanje visoke točnosti uz minimalnu računalnu složenost. Ova arhitektura koristi dubinski odvojive konvolucije (engl. *Depthwise Separable Convolutions*) kako bi smanjila broj parametara i povećala učinkovitost, što je čini idealnim rješenjem za uređaje s ograničenim resursima. U kontekstu klasifikacije hrane, MobileNetV3 se može podesiti za precizno prepoznavanje različitih vrsta hrane i procjenu njihove nutritivne vrijednosti [30].

NASNet je model razvijen primjenom metode *Neural Architecture Search* (NAS), koja koristi algoritme za automatsko istraživanje optimalnih arhitektura neuronskih mreža [31]. Ovaj model postiže visoku preciznost u klasifikaciji slika i često se koristi u zadacima gdje je potrebno postići ravnotežu između točnosti i računalne složenosti. Primjena NASNet-a u klasifikaciji hrane omogućava precizno prepoznavanje i procjenu hrane, čak i u složenim situacijama [31].

### **3.5. Pregled modela za estimaciju kalorija**

Sljedeća korištena arhitektura za estimaciju kalorija nije primarno razvijena za potrebe ovog diplomskog rada te je iskorištena već postojeća arhitektura na kojoj se evaluiraju standardizirani skupovi podataka. Sama standardizacija je osmišljena kako bi se prilagodila formatima arhitektura korištenih za estimaciju kalorija. Kako se njen značaj ne bi zanemario u nastavku je površinski opisana arhitektura i njene glavne značajke. U središtu arhitekture nalaze se unaprijed trenirani CNN modeli, poput MobileNetV3 i NASNetMobile, koji su temeljni blokovi za ekstrakciju značajki. Ovi modeli su inicijalno trenirani na velikom skupu podataka čime su već razvili sposobnost prepoznavanja općih vizualnih obrazaca u slikama.

#### **3.5.1. Korištene biblioteke**

TensorFlow je glavna biblioteka korištena za implementaciju i treniranje modela strojnog učenja u ovom radu. Razvijen od strane Googlea, TensorFlow je platforma otvorenog koda za izgradnju i treniranje modela dubokog učenja, s posebnim naglaskom na rad s konvolucijskim neuronskim mrežama (CNN) i obradom velikih skupova podataka. [44] Unutar ovog rada, TensorFlow je korišten za definiranje i treniranje modela koji predviđaju kategorije hrane, sastojke i nutritivne

vrijednosti. Keras je API unutar TensorFlow platforme koji pojednostavljuje proces izgradnje modela. Ovaj alat omogućava brzu izradu i testiranje modela neuronskih mreža putem jednostavnih i intuitivnih funkcija. [45] Unutar ovog rada, Keras se koristi za definiranje slojeva modela i za upravljanje složenim mrežnim arhitekturama putem pozivanja jednostavnih funkcija.

TensorFlow Datasets (TFDS) je modul unutar TensorFlow platforme koji omogućuje rad sa skupovima podataka [46]. U ovom radu, korišten je za učitavanje i obradu više velikih skupova podataka, uključujući slike hrane i pridružene meta podatke. Funkcionalnosti TFDS-a omogućuju automatsko preuzimanje, predobradu i pretvorbu podataka u formate prikladne za treniranje modela.

Google Colab (*drive module*) je korišten za povezivanje i pohranu podataka na Google Drive platformi, omogućujući pristup velikim skupovima podataka tijekom treniranja modela.

Ipybn je korišten za integraciju s Jupyter Notebook formatom i omogućuje rad s funkcijama unutar bilježnica u drugim okruženjima.

### **3.5.2. Aktivacijske funkcije, hiperparametri i optimizator**

U različitim slojevima modela primjenjuju se različite aktivacijske funkcije koje transformiraju izlaze slojeva kako bi se prilagodili zadacima kao što su klasifikacija i regresija. Glavna aktivacijska funkcija koja se koristi u konvolucijskim i potpuno povezanim slojevima je ReLU, koja omogućava modelu da uči nelinearne odnose u podacima, što je ključno za prepoznavanje složenih uzoraka u slikama. ReLU postavlja sve negativne vrijednosti na nulu, čime se ubrzava konvergencija tijekom procesa treniranja [47]. Za završni sloj u zadatku klasifikacije kategorija koristi se Softmax. Ova funkcija pretvara izlaze modela u vjerojatnosti za svaku kategoriju, pri čemu kategorija s najvišom vjerojatnošću predstavlja konačnu predikciju [35].

Sigmoid funkcija se koristi za višestruku klasifikaciju, pretvarajući izlaze u vrijednosti između nula i jedan, što omogućava modelu da neovisno procijeni prisutnost ili odsutnost svake jedinice [48]. Za regresijske zadatke, poput predikcije kontinuiranih vrijednosti, koristi se linearna aktivacijska funkcija. Svaki predviđani parametar ima zasebni linearni izlaz, čime se omogućava precizna predikcija svih vrijednosti [49].

Hiperparametri su ključni za upravljanje procesom treniranja modela i njegovom konačnom izvedbom. U nastavku su predstavljeni hiperparametri korišteni u ovom diplomskom radu.

Veličina *batch-a* određuje broj uzoraka koji se koriste za ažuriranje težina modela tijekom svake iteracije. Veća veličina *batcha* omogućuje stabilnije treniranje, dok manja veličina omogućuje bržu prilagodbu težina [50]. Stopa učenja određuje brzinu kojom se model prilagođava promjenama u težinama tijekom treniranja. Manja stopa učenja osigurava stabilnije treniranje, ali može zahtijevati više iteracija, dok veća stopa može ubrzati proces, ali i uzrokovati nestabilnost [51]. Broj epoha određuje koliko puta će model proći kroz cijeli skup podataka tijekom treniranja. Više epoha omogućuje modelu da bolje nauči podatke, ali postoji rizik od pretreniranosti ako se model trenira predugo [52].

Pored navedenih hiperparametara, korištene su i određene *callback funkcije* za optimizaciju procesa treniranja. Rano zaustavljanje je funkcija prekida treniranje kada ne dođe do poboljšanja u `val_loss`-u tokom određenog broja epohi, čime se sprječava pretjerano treniranje modela [54]. `ReduceLROnPlateau` je funkcija smanjuje stopu učenja ako `val_loss` ostane nepromijenjen, smanjujući je odabranim faktorom do odabrane minimalne vrijednosti. Ovaj mehanizam omogućava preciznije prilagodbe tokom treniranja i sprječava nagle promjene u prilagodbi težina [35].

Optimizator koji je korišten za obuku modela je Adam (engl. *Adaptive Moment Estimation*), jedan od najkorištenijih optimizatora u dubokom učenju. Adam spaja prednosti tradicionalnih optimizatora kao što su Stochastic Gradient Descent (SGD) i RMSprop, omogućavajući bržu konvergenciju i efikasnije upravljanje varijabilnošću tokom procesa treniranja. Ovaj optimizator automatski prilagođava stope učenja za svaku težinu modela, što olakšava obuku složenih modela poput ovog [53].

### **3.5.3. Pregled arhitekture i slojeva modela**

Arhitektura korištena u ovom radu oslanja se na visezadaćni model dubinskog učenja koji je dizajniran za rješavanje više zadataka različitih tipova istovremeno, kao što su npr. klasifikacija i regresija. Model se sastoji od dvije ključne komponente: konvolucijskih slojeva za obradu slikovnih podataka i potpuno povezanih slojeva za generiranje predviđanja.

Jedna od ključnih komponenti ovog modela je primjena konvolucijskih neuronskih mreža (CNN). CNN-ovi su izuzetno važni za analizu slika hrane, koja je glavni fokus ovog istraživanja.



Proces ekstrakcije značajki započinje primjenom niza konvolucijskih filtera (kernela) koji prolaze kroz sliku i prepoznaju ključne vizualne elemente. Npr. donji slojevi modela mogu detektirati osnovne komponente poput rubova, dok viši slojevi prepoznaju složenije obrasce i strukture, kao što su oblici hrane ili detalji sastojaka. Ovi slojevi djeluju u određenom redoslijedu, pri čemu se donji slojevi usredotočuju na jednostavnije značajke, dok viši slojevi postaju specijalizirani za prepoznavanje kompleksnijih vizualnih karakteristika.

U konvolucijskim slojevima, *max-pooling* slojevi služe za smanjenje dimenzionalnosti podataka. Da bi se osigurala stabilnost tijekom treniranja modela, primjenjuje se normalizacija podataka. Konvolucijski slojevi u ovoj arhitekturi potječu iz unaprijed istreniranih modela, što također određuje broj slojeva unutar tog segmenta mreže te nisu pojedinačno analizirani.

Nakon što CNN analizira ulazne slike i izvuče ključne značajke, ti se podaci prosljeđuju u potpuno povezane slojeve. U ovoj fazi obrade, model donosi predikcije za svaki od zadanih zadataka. U toj fazi, arhitektura modela se dijeli na tri paralelne grane, pri čemu svaka grana rješava jedan od specifičnih zadataka:

- Dijeljeni slojevi: Jedan potpuno povezani sloj s 2048 neurona koji prenosi zajedničke značajke u tri odvojene grane.
- Klasifikacija hrane: Ova grana koristi jedan sloj s 512 neurona koji predviđa kojoj kategoriji hrana pripada. Aktivacijska funkcija je Softmax koja omogućava modelu da odredi kojoj kategoriji hrana pripada, pri čemu su sve kategorije međusobno isključuju, a zbroj vjerojatnosti svih kategorija iznosi 1.
- Klasifikacija sastojaka: Ova grana koristi tri sloja s 1024, 512 i 256 neurona. Svaki sloj pomaže u predikciji prisutnosti ili odsutnosti sastojaka. Aktivacijska funkcija je Sigmoid koja omogućava modelu da samostalno predviđa prisutnost ili odsutnost svakog sastojka, s obzirom na to da jedan uzorak može pripadati više različitih klasam. Svaki sastojak se neovisno ocjenjuje na osnovu vrijednosti između 0 i 1, što označava njegovu prisutnost ili odsutnost.
- Regresija nutritivnih vrijednosti: Svaki od četiri nutritivna parametra (kalorije, proteini, ugljikohidrati, masti) koristi dva sloja s 64 i 32 neurona za generiranje kontinuiranih predikcija. Ovi slojevi nemaju aktivacijsku funkciju jer je cilj generirati numeričke vrijednosti.

Budući da se ove dvije arhitekture koriste za usporedbu, primijenjeni su identični hiperparametri. Također, optimizator koji se koristi u obje arhitekture je Adam. Hiperparametri za obje arhitekture prikazani su u Tablici 3.1.

Tablica 3.1. Korišteni hiperparametri za sve konfiguracije modela

Hiperparametar	Vrijednost
Veličina batcha	256
Broj epoha	5
Stopa učenja	0.001
ReduceLROnPlateau	0.5
Težine gubitka	1.0
Rano zaustavljanje	Nakon 8 epoha

## 4. Provedeni eksperiment i dobiveni rezultati

Prilikom odabira okruženja za izvođenje eksperimenta razmatrane su različite opcije. Kao glavne su se isticala lokalno okruženje laptopa, Anaconda environment te online virtualno okruženje Google Colab u kombinaciji s Google Drive-om. Budući da je od važnosti imati mogućnost pohrane veće količine podataka i nesmetan rad s algoritmima bez oslanjanja na performanse laptopa, koristi se Google Colab kao optimalno rješenje. U ovom radu fokus je na implementaciji cjevovodnog sustava, što je zahtijeva dodatne modifikacije kako bi cijeli kod funkcionirao ispravno. Eksperimenti su osmišljeni da testiraju standardizirane skupove podataka na jednom od dva razvijena modela za procjenu kalorija, a rezultati se zatim uspoređuju. Prikazani su rezultati dobiveni na pojedinačnim skupovima podataka, zatim rezultati na različitim kombinacijama skupova podataka, gdje se spajaju dva ili više skupova podataka i na kraju su kombinirani svi skupovi podataka kako bi se dobili konačni rezultati. Istražuje se pretpostavka može li ovako strukturiran skup podataka dovesti do lošijih ili boljih rezultata. Naravno, pretpostavka je da su mnogi koraci izvedeni neadekvatno, što će utjecati na kvalitetu. Međutim, naglasak je na tome koliko kvalitetan rezultat pruža konačno sastavljen skup podataka iz različitih izvora u odnosu na pojedinačne skupove ili njihove kombinacije.

Tablica 4.1. Korištena konfiguracija

Google Drive pro 2TB
Google collab Free:
Ukupni RAM: 12GGB
Ukupni prostor na disku: ['108GB']
Procesor: Intel(R) Xeon(R) CPU @ 2.20GHz
Python verzija: 3.10.12

## 4.2. Odabrane metode evaluacije

Za problem klasifikacije u praksi najčešće korištene metrike uključuju točnost (engl. *accuracy*), preciznost (engl. *precision*), odziv (engl. *recall*) i F1-skor prema članku [33]. Točnost je osnovna mjera koja pokazuje postotak ispravno klasificiranih instanci u odnosu na ukupan broj instanci. Ova mjera je korisna kada su podaci uravnoteženi, ali može biti varljiva kod neravnoteže u broju klasa. Preciznost mjeri točnost pozitivnih predikcija, odnosno koliko često model ispravno identificira pozitivne instance u odnosu na sve instance predviđene kao pozitivne. Odziv s druge strane, mjeri koliko uspješno model identificira sve stvarne pozitivne instance. Za zadatke regresije, gdje model predviđa kontinuirane vrijednosti poput kalorija, proteina, ugljikohidrata i masti, najčešće korištene metrike su srednja apsolutna pogreška (MAE) i srednja kvadratna pogreška (MSE) [34]. Srednja apsolutna pogreška (MAE) mjeri prosječnu apsolutnu razliku između predviđenih i stvarnih vrijednosti. MAE je intuitivna i daje jasnu sliku o pogreškama u istim jedinicama kao i ciljna varijabla, što je čini korisnom za praktične primjene. Srednja kvadratna pogreška (MSE) mjeri prosječnu kvadratnu pogrešku između predviđenih i stvarnih vrijednosti. Budući da kvadrira pogreške, MSE daje veću težinu većim pogreškama, što je korisno kada su velike pogreške neprihvatljive u modelu.

U ovom diplomskom radu, gdje se razvija model koji obavlja više zadataka klasifikacije i regresije, koristi se kombinacija različitih metrika kako bi se precizno procijenile performanse modela. Za klasifikaciju kategorija hrane koristi se točnost, preciznost i odziv, ključne metrike za ocjenu uspješnosti modela u prepoznavanju različitih kategorija. Važno je da model ne samo pravilno identificira većinu kategorija, već i da minimizira broj lažno pozitivnih i lažno negativnih predikcija, što ove metrike omogućuju [34]. Za predviđanje nutritivnih vrijednosti, poput kalorija, proteina, masti i ugljikohidrata, odabrana je srednja apsolutna pogreška (MAE) zbog njene jednostavnosti i lakoće interpretacije. MAE omogućuje jasno razumijevanje prosječne pogreške modela u istim jedinicama kao i ciljane varijable [35]. Preciznost i odziv još se koriste za višestruku klasifikaciju sastojaka unutar jela, što je od ključne važnosti za točne nutritivne analize. Ove metrike pomažu u procjeni sposobnosti modela da ispravno identificira sve relevantne sastojke, čime se dodatno osigurava preciznost u krajnjim predikcijama.

## 4.2. Rezultati za pojedinačne skupove podataka

U nastavku su prikazani rezultati izvršavanja dva algoritma. Prvi algoritam koji se primjenjuje je MobileNetV3, a zatim NASNetModel. Oba modela su primijenjena na svakom skupu podataka uz korištenje iste arhitekture i istog broja epoha.

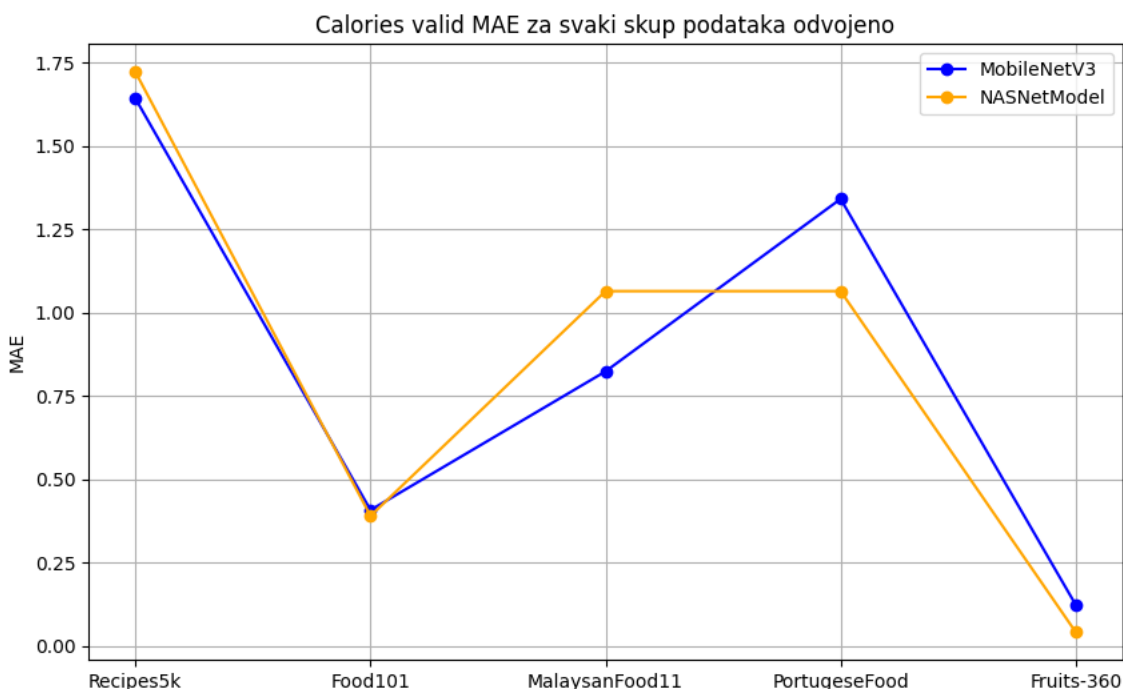
Rezultati u nastavku pružaju uvid u performanse svakog skupa podataka odvojeno, što se kasnije iskoristilo za usporedbi performansi različitih algoritama na istim skupovima podataka te na kombinacijama skupova podataka.

Tablica 4.2. Usporedba rezultata za pojedinačni skup podataka

Skup podataka	Recipes5k	Food101	MalasyanFood11	PortugeseFood	Fruits-360
Broj jedinki za treniranje	4826	19810	6975	6221	7590
Carbs valid <b>MAE</b>	0.1767 0.2613	0.0653 0.0686	0.1155 0.1159	0.1623 0.1027	0.1234 0.1105
Protein valid <b>MAE</b>	0.1813 0.1317	0.0468 0.0367	0.1708 0.1634	0.1115 0.0670	0.0620 0.0610
Fat valid <b>MAE</b>	0.2242 0.1879	0.0696 0.0666	0.1001 0.0998	0.1105 0.1495	0.0975 0.0522
Category valid <b>ACC</b>	0.4767 0.2746	0.7494 0.6285	0.9068 0.7907	0.9502 0.8762	0.9993 0.9545
Ingredients valid <b>Precision</b>	0.4126 0.4818	0.8994 0.8489	0.9022 0.8266	0.8825 0.8856	0.8526 1.0000
Calories valid <b>MAE</b>	1.6417 1.7213	0.4082 0.3906	0.8239 0.6607	1.3409 1.0644	0.1244 0.0425

Oba modela, MobileNetV3 i NASNetModel, pokazuju slične performanse u procjeni kalorija na većini skupova podataka, no postoje određene varijacije koje je važno primijetiti. Najveća pogreška u predikciji kalorija uočava se na skupu podataka Recipes5k, gdje oba modela postižu MAE iznad 1.5, s tim da je NASNetModel blago precizniji od MobileNetV3. Slično je i na skupu PortugueseFood, gdje MobileNetV3 ima veću pogrešku u odnosu na NASNetModel.

S druge strane, na skupu Food101 oba modela pokazuju gotovo identičnu točnost, s MAE ispod 0.5. Skupovi podataka MalaysianFood11 i Fruits-360 pokazuju niske pogreške s MAE ispod 1.0, ali s nešto većim razlikama između modela. MobileNetV3 ima bolji rezultat na MalaysianFood11, dok NASNetModel pokazuje precizniju predikciju na skupu Fruits-360. Razlike između ta dva modela su na nekim mjestima potpuno neznačajne, dok su na nekim mjestima ipak vidljivije te je vrlo teško bilo favorizirati ijedan od ta dva modela. Što se tiče samih skupova podataka, vidi se jasna razlika u varijaciji između estimacije kalorija. Primjerice, kompleksniji skupovi podataka koji sadrže kompleksna jela imaju loše rezultate, dok jednostavniji, poput Fruits360, gdje se određuju kalorije samo određenog voća, imaju odlične rezultate. Sve je to popraćeno grafikonom na Slici 4.1.

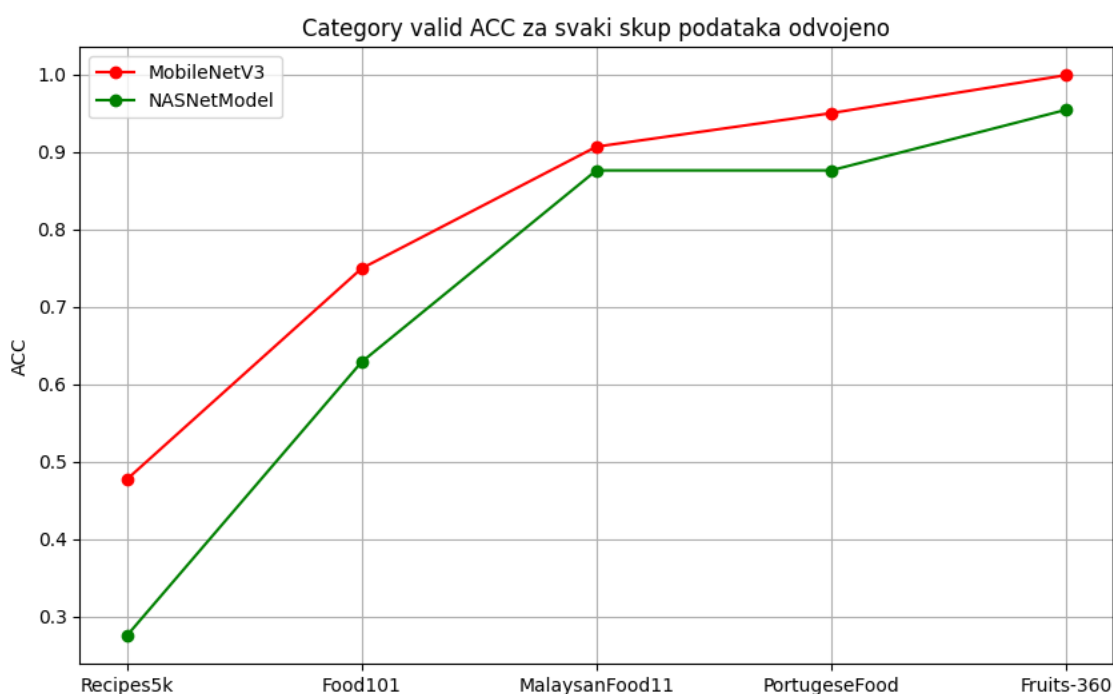


Slika 4.1. Prikaz estimacije kalorija za pojedinačni skup podataka metrikom MAE

Iz grafikona sa Slike 4.2. je jasno da MobileNetV3 dosljedno postiže bolje rezultate u odnosu na NASNetModel u zadatku kategorizacije hrane. MobileNetV3 pokazuje stabilno visoku točnost, osobito na skupovima MalaysianFood11, PortugeseFood i Fruits-360, gdje točnost

dostiže gotovo 1.0, što sugerira izvrsne performanse u raznolikim i složenim zadacima klasifikacije.

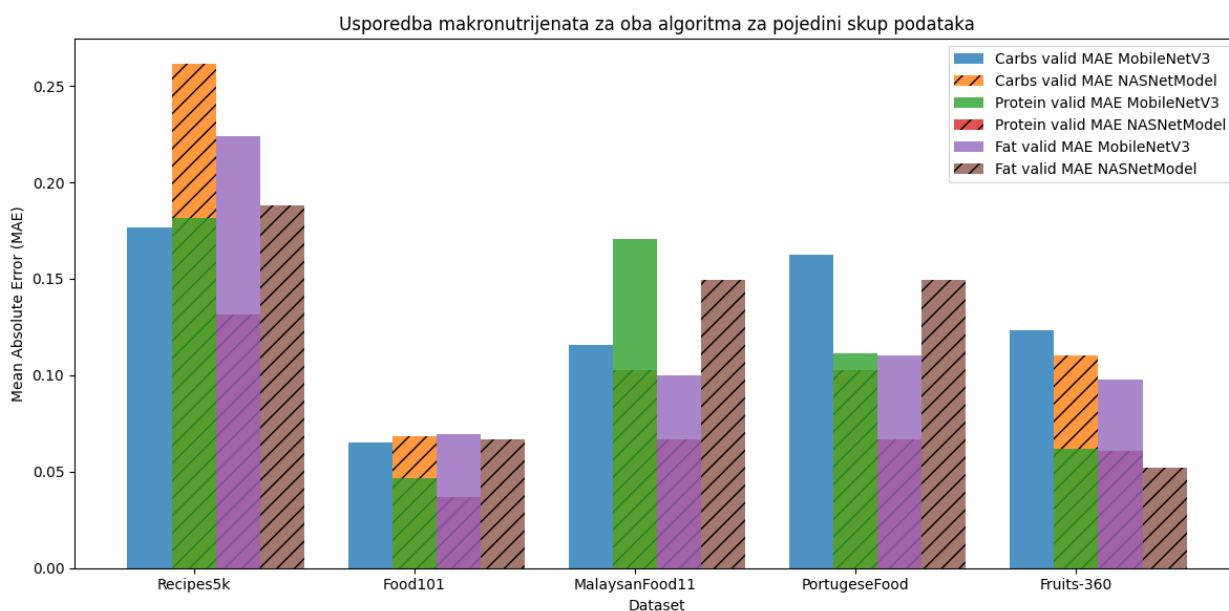
S druge strane, NASNetModel započinje s nižom točnošću na Recipes5k i Food101, s vrijednostima između 0.3 i 0.6, ali postepeno poboljšava performanse na MalaysianFood11 i PortugueseFood. Iako se performanse NASNetModel-a značajno poboljšavaju na kasnijim skupovima, MobileNetV3 i dalje pokazuje superiornost u ukupnoj točnosti, osobito na složenijim skupovima podataka. Ovaj rezultat sugerira da je MobileNetV3 bolji izbor za zadatke klasifikacije, osobito kada se radi o skupovima podataka koji zahtijevaju veću preciznost i točnost.



Slika 4.2. Prikaz rezultata klasifikacije pojedinih skupova podataka za oba algoritma

Grafikon prikazan na Slici 4.3. prikazuje usporedbu srednje apsolutne pogreške (MAE) za predikcije triju makronutrijenata (ugljikohidrati, proteini i masti) korištenjem dva različita modela, MobileNetV3 i NASNetModel. Na skupu Recipes5k, model NASNetModel pokazuje znatno veće pogreške u predikciji proteina i masti u usporedbi s model MobileNetV3, dok model MobileNetV3 daje stabilnije i preciznije rezultate za sve makronutrijente. Na skupu Food101, NASNetModel postiže nešto bolje rezultate od

MobileNetV3, s manjim MAE-om za sve makronutrijente. Kod MalaysianFood11 primjetan je obrnuti trend, gdje MobileNetV3 daje bolje rezultate za proteine i masti, dok NASNetModel postiže bolju točnost za ugljikohidrate. Na skupu PortugueseFood oba modela pokazuju slične rezultate, no MobileNetV3 je nešto precizniji u predikciji ugljikohidrata, dok NASNetModel pokazuje nešto bolje rezultate za proteine. Skup Fruits-360 pokazuje uravnotežene performanse između dva modela, gdje su rezultati bliski, ali MobileNetV3 ima blagu prednost u predikciji ugljikohidrata i masti.



Slika 4.3. Prikaz rezultata klasifikacije nutritivnih vrijednosti za oba algoritma

### 4.3. Rezultati Kombinacijom dvaju skupova podataka

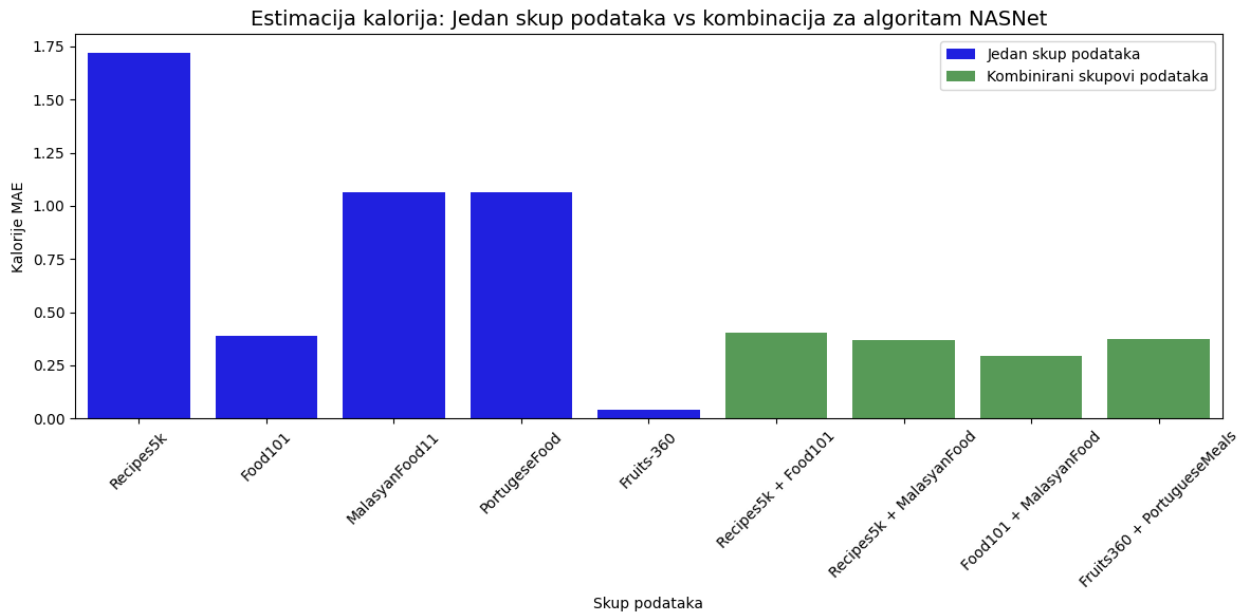
U nastavku su prikazani rezultati izvršavanja za jedan algoritam, NASNet. Broj epoha ostao je nepromijenjen, a primijenjene su različite kombinacije skupova podataka u kojima se dva skupa kombiniraju u jedan. Performanse ovih kombinacija uspoređene su s performansama pojedinačnih skupova podataka.



Tablica 4.3. Usporedba rezultata za dvije kombinacije skupova podataka

<b>Skup podataka</b>	Recipes5k + Food101	Recipes5k + MalasyanFood	Food101 + MalasyanFood	Fruits360 + PorugueseMeals
Broj jedinki za treniranje	24636	11801	26785	13811
Carbs valid <b>MAE</b>	0.0613	0.0833	0.0422	0.0828
Protein valid <b>MAE</b>	0.0335	0.0596	0.0260	0.0449
Fat valid <b>MAE</b>	0.0486	0.0527	0.0427	0.0528
Category valid <b>ACC</b>	0.6365	0.8195	0.6881	0.8896
Ingredients valid <b>Precision</b>	0.8231	0.9349	0.7087	0.9641
Calories valid <b>MAE</b>	0.4027	0.3668	0.2965	0.3731

Slika 4.4. prikazuje usporedbu srednje apsolutne pogreške (MAE) u procjeni kalorija za algoritam NASNet koristeći dva pristupa: treniranje na pojedinačnim skupovima podataka i treniranje na kombiniranim skupovima. Plavi stupci predstavljaju rezultate kada je model treniran na jednom skupu podataka, dok zeleni stupci prikazuju rezultate kada se koristi kombinacija više skupova. Iz grafikona je vidljivo da su pogreške značajno manje kada se trenira na kombiniranim skupovima podataka, posebice za MalaysianFood11 i Fruits-360. Npr. treniranje na pojedinačnim skupovima kao što su Recipes5k i PortugueseFood rezultira većim pogreškama, s MAE iznad 1.5, dok kombinacija skupova, kao što su Food101 i MalaysianFood, smanjuje pogreške na ispod 0.5 MAE. Ovi rezultati sugeriraju da kombiniranje različitih skupova podataka omogućuje modelu bolji obuhvat raznolikosti u nutritivnim karakteristikama hrane, što rezultira boljom preciznošću u procjeni kalorija.



Slika 4.4. Usporedba rezultata estimacije kalorija za jedan skup podataka sa kombinacijom dvaju skupova podataka

#### 4.4. Rezultati Kombinacijom triju skupova podataka

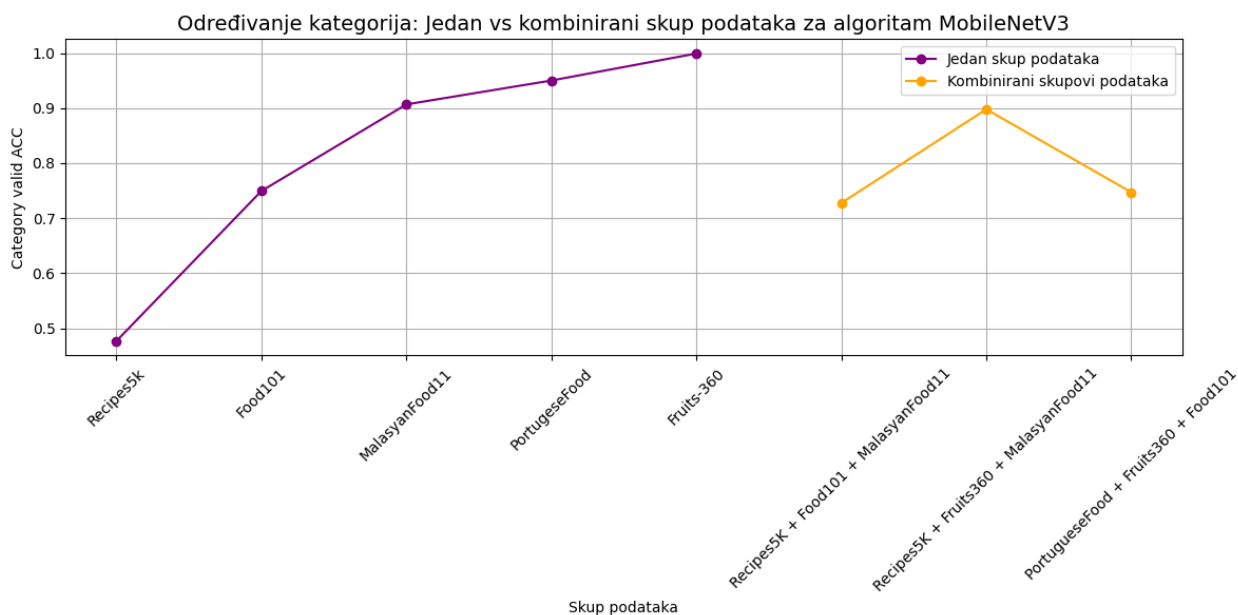
Prikazani rezultati odnose se na izvršavanje algoritma MobileNetv3. Broj epoha ostao je konstantan, dok su isprobane različite kombinacije podataka, pri čemu su tri različita skupa spojena u jedan. Performanse tih kombinacija analizirane su u usporedbi s performansama pojedinačnih skupova podataka.

Tablica 4.4. Usporedba rezultata za tri kombinacije skupova podataka

<b>Skup podataka</b>	Recipes5k + Food101+ MalasyanFood11	Recipes5k + Fruits360 + MalasyanFood11	PortugueseFood + Fruits360 + Food101
Broj jedinki za treniranje	31611	19391	33621
Carbs valid <b>MAE</b>	0.0507	0.0414	0.0800
Protein valid <b>MAE</b>	0.0373	0.0533	0.0333
Fat valid <b>MAE</b>	0.0560	0.0357	0.0548
Category valid <b>ACC</b>	0.7276	0.8979	0.7472
Ingredients valid <b>Precision</b>	0.6316	0.8862	0.8862
Calories valid <b>MAE</b>	0.3595	0.2258	0.3854

Iz grafikona prikazanog na Slici 4.5. vidljivo je da MobileNetV3 postiže vrlo visoku točnost kada se trenira na pojedinačnim skupovima podataka, posebice na MalaysianFood11 i Fruits-360, gdje točnost dostiže gotovo 1.0. Ovo pokazuje da model vrlo dobro funkcionira kada je treniran na specifičnim, dobro definiranim skupovima podataka. Međutim, kada se koriste kombinirani skupovi podataka, točnost i dalje ostaje visoka, ali ne dostiže razine kao kod treniranja na pojedinačnim skupovima. Najviša točnost među kombiniranim skupovima postiže se za kombinaciju Fruits360 i Food101, ali je i dalje nešto niža u usporedbi s performansama na pojedinačnim skupovima MalaysianFood11 ili Fruits-360.

Ovaj rezultat sugerira da iako kombinacija podataka može poboljšati performanse modela u smislu generalizacije, MobileNetV3 pokazuje bolje rezultate u specifičnim zadacima kada je treniran na pojedinačnim skupovima podataka. U specifičnim situacijama gdje su podaci homogeni ili usmjereni na određenu vrstu hrane, treniranje na jednom skupu omogućava postizanje vrlo visoke točnosti. Kombinacija podataka pruža dodatnu generalizaciju, ali točnost može biti blago kompromitirana zbog šire varijacije u podacima



Slika 4.5. Usporedba rezultata određivanja kategorija za pojedini skup podataka sa kombinacijom triju skupova podataka

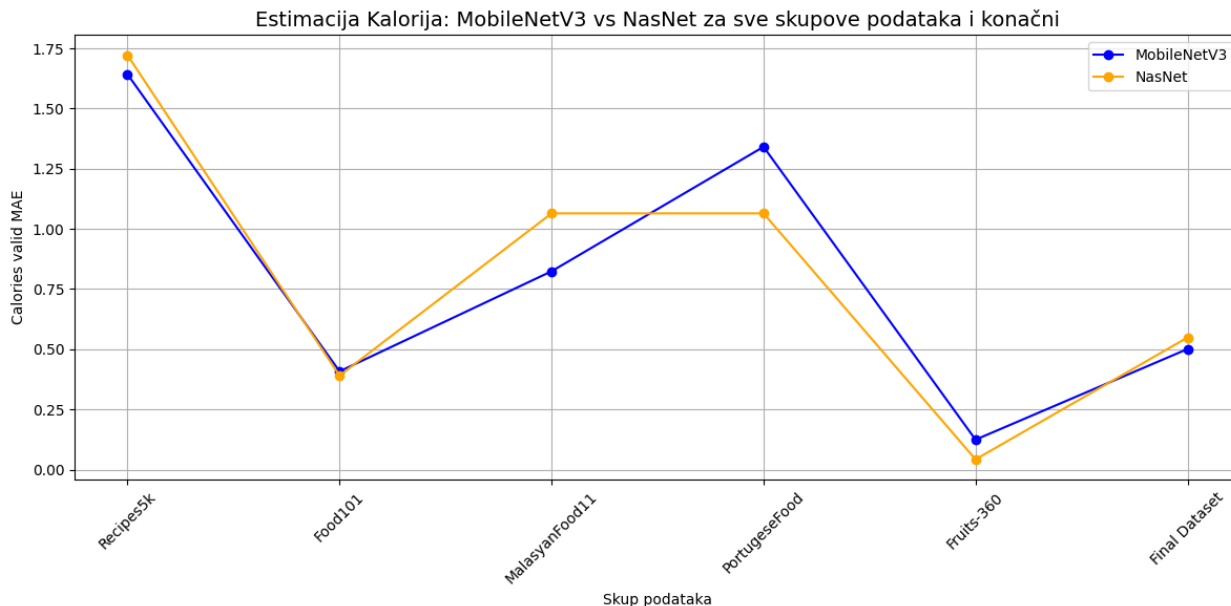
## 4.5. Rezultati za cijeli skup podataka

Prikazani su rezultati izvršavanja oba algoritma, NASNet i MobileNetv3. Broj epoha, kao i arhitektura modela, ostali su nepromijenjeni tijekom svakog izvođenja. Evaluacija je provedena na konačnom skupu podataka koji objedinjuje sve prethodno korištene i opisane skupove podataka, što je bio glavni cilj ovog rada. Ova tablica omogućava interpretaciju rezultata korištenjem standardiziranog konačnog skupa podataka u odnosu na pojedinačne skupove. Mogu se donijeti i zaključci o tome je li ovaj pristup doprinio poboljšanju procjene kalorija.

Tablica 4.5. Rezultati sviju zadataka na oba algoritma koristeći finalni skup podataka

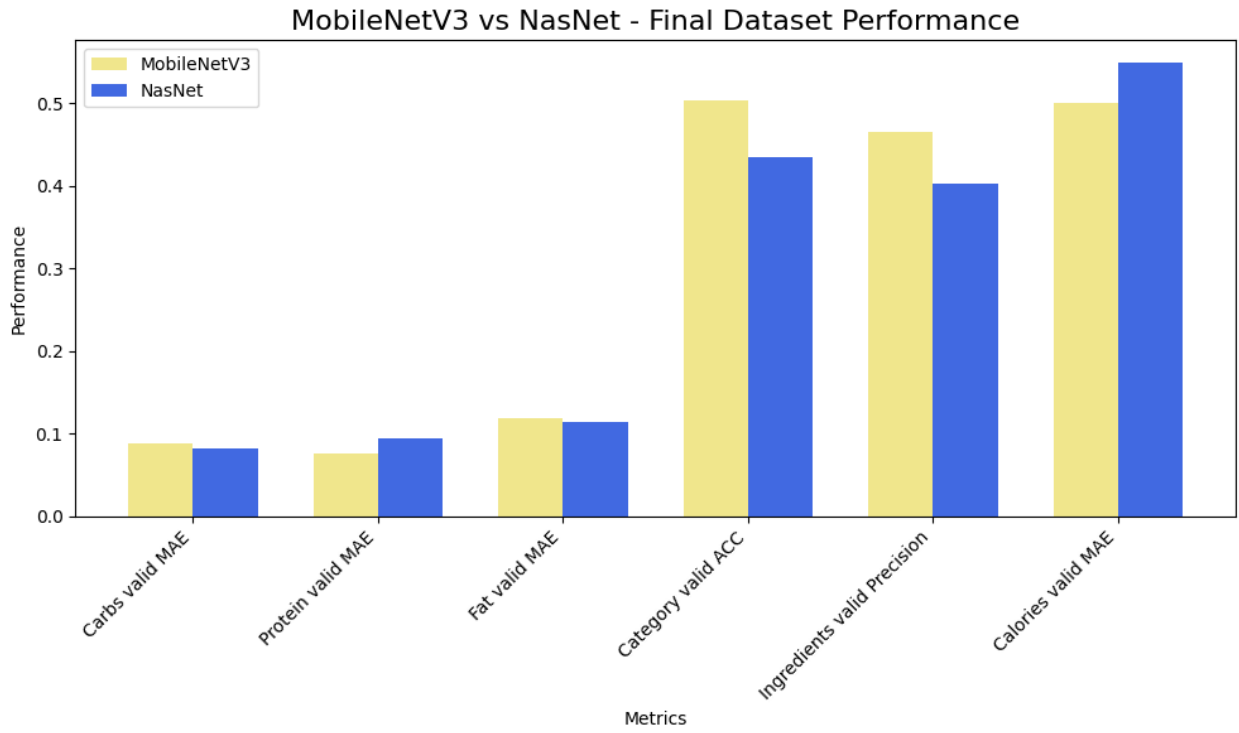
<b>Algoritam</b>	<b>MobileNetV3</b>	<b>NasNET</b>
Broj jedinki za treniranje	45422	45422
Carbs valid <b>MAE</b>	0.0881	0.0822
Protein valid <b>MAE</b>	0.0763	0.0937
Fat valid <b>MAE</b>	0.1192	0.1146
Category valid <b>ACC</b>	0.5032	0.4343
Ingredients valid <b>Precision</b>	0.4653	0.4032
Calories valid <b>MAE</b>	0.5010	0.5488

Iz grafikona prikazanog na Slici 4.6. je vidljivo da "Final Dataset" prikazuje znatno nižu srednju apsolutnu pogrešku u odnosu na pojedinačne skupove podataka, što naglašava ključnu prednost korištenja većeg, kombiniranog skupa podataka. Skupovi poput Recipes5k i PortugueseFood pokazuju visoke pogreške (iznad 1.5), dok je MAE na "Final Datasetu" znatno niži, što sugerira da treniranje modela na većem skupu podataka omogućuje bolju generalizaciju i preciznije predikcije kalorija. Ova niža pogreška rezultat je šireg obuhvata različitih vrsta hrane i nutritivnih profila, čime model može bolje naučiti i prepoznati raznovrsne obrasce u podacima. Konačna MAE vrijednost za "Final Dataset" vrlo je bliska optimalnim performansama koje modeli postižu na manjim i specifičnim skupovima, poput Fruits-360, gdje je pogreška također vrlo niska. To dodatno potvrđuje da kombinacija više skupova podataka ne samo da omogućava bolju generalizaciju, nego i značajno smanjuje pogrešku.



Slika 4.6. Estimacija kalorija: MobileNetV3 vs NasNet za sve skupove podataka

Iz grafikona prikazanog na Slici 4.7. vidljivo je da su performanse MobileNetV3 i NASNet vrlo slične kada je riječ o metrike vezane uz makro nutrijente (ugljikohidrati, proteini i masti). Za sve makro nutrijente, srednja apsolutna pogreška (MAE) je gotovo jednaka za oba modela, što sugerira da su oba modela podjednako sposobna za predviđanje nutritivnih elemenata. Međutim, MobileNetV3 značajno nadmašuje NASNet u zadacima koji se odnose na klasifikaciju kategorija i prepoznavanje sastojaka. Točnost kategorija za MobileNetV3 postiže 0.5, dok NASNet postiže nešto niže vrijednosti. Slično, MobileNetV3 također nadmašuje NASNet u preciznosti sastojaka (engl. *Ingredients output Precision*), što ukazuje na to da je MobileNetV3 bolji u preciznom identificiranju nutritivnih svojstava i klasifikaciji hrane. Za procjenu kalorija, oba algoritma postižu rezultate iznad 0.5, međutim NASNet pokazuje nešto bolje rezultate te bi se moglo reći da je bolji za problem estimacije kalorija.



Slika 4.7. MobileNetV3 vs NasNet konačne performanse

## 5. Zaključak

Analiza relevantne literature obuhvatila je složenost procjene kalorija na temelju slika hrane, pri čemu je posebno izazovna procjena volumena hrane. Raznolikost podataka iz različitih izvora prepoznata je kao još jedan od ključnih izazova, te predstavlja početni korak u koracima procesa estimacije kalorija. Skupovi podataka za prepoznavanje hrane dolaze iz različitih regija i često imaju neujednačene strukture, što otežava njihovu integraciju u jedinstveni sustav za treniranje modela. Analizirani su i odabrani javno dostupni skupovi podataka, poput Food101 i Fruits360, koji su se pokazali korisnima za specifične zadatke, ali nisu imali kompletne nutritivne podatke potrebne za preciznu estimaciju kalorija. Unatoč brojnim izazovima, postignuta je djelomična automatizacija procesa standardizacije podataka koja uključuje učitavače podataka, univerzalni onehotencoder te optimizirani format za spremanje podataka - tfrecord. Identificirana su dva glavna tipa skupova podataka: oni koji koriste mape za kategorizaciju i oni koji sadrže meta datoteke. Ova automatizacija omogućila je ubranu obradu podataka, ali nije u potpunosti riješila problem zbog velike raznolikosti podataka i kompleksnosti potrebne obrade. Evaluacija modela, u kojoj su uspoređeni MobileNetV3 i NASNet, pokazala je da MobileNetV3 ostvaruje bolje rezultate u zadacima klasifikacije hrane i prepoznavanja sastojaka, što ga čini pogodnijim za zadatke koji zahtijevaju precizno kategoriziranje hrane. S druge strane, NASNet je postigao nešto bolje rezultate u procjeni kalorija, što ukazuje na njegovu sposobnost da bolje procijeni nutritivnu vrijednost jela. Ipak, zbog dosljedno bolje performanse u većini zadataka, MobileNetV3 se pokazao kao optimalniji izbor za većinu scenarija u okviru ovog istraživanja. Uočeno je da skupovi podataka koji su manji i specifični, poput MalaysianFood11, omogućuju brže treniranje modela i preciznije rezultate zbog homogenosti podataka. Kombinacijom različitih skupova podataka postignute su bolje performanse, jer je model imao priliku trenirati na raznovrsnijim slikama hrane, čime se poboljšala ukupna točnost i smanjila srednja apsolutna pogreška (MAE) u estimaciji kalorija. Veliki skupovi podataka, poput "Final Dataset", pokazali su se ključnima za postizanje bolje generalizacije modela. Upotreba velikog skupa podataka omogućila je precizniju estimaciju kalorija i nutritivnih vrijednosti u složenijim situacijama, jer pružaju širi spektar različitih tipova hrane. Ovim rezultatima se potiče daljnje istraživanje standardizacije skupova podataka te postizanje boljih rezultata.



## Literatura

- [1] F. P. W. Lo, Y. Sun, J. Qiu and B. Lo, "Image-Based Food Classification and Volume Estimation for Dietary Assessment: A Review," in *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 1926-1939, July 2020, doi: 10.1109/JBHI.2020.2987943
- [2] Gilal, N.U., Al-Thelaya, K., Al-Saeed, J.K. *et al.* Evaluating machine learning technologies for food computing from a data set perspective. *Multimed Tools Appl* **83**, 32041–32068 (2024). <https://doi.org/10.1007/s11042-023-16513-4>
- [3] Image Based Food Calories Estimation Using Various Models of Machine Learning , Haoyu Hu;Zihao Zhang;Yulin Song, 2020
- [4] Refined Image Segmentation for Calorie Estimation of Multiple-Dish Food Items, Parth Poply; J. Angel,Arul Jothi, 2021
- [5] Instance-Aware Semantic Segmentation for Food Calorie Estimation using Mask R-CNN Reza Dea Yogaswara;Eko Mulyanto Yuniarno;Adhi Dharma Wibawa 2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)
- [6] Image-Based Estimation of Real Food Size for Accurate Food Calorie Estimation Takumi Ege;Yoshikazu Ando;Ryosuke Tanno;Wataru Shimoda;Keiji Yanai2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)
- [7] A New Large-scale Food Image Segmentation Dataset and Its Application to Food Calorie Estimation Based on Grains of RiceEge, T; Shimoda, W and Yanai, K 5th International Workshop on Multimedia Assisted Dietary Management (MADiMa) 2019
- [8] A Hybrid Network Based on GAN and CNN for Food Segmentation and Calorie EstimationR. Jaswanthi;E. Amruthatulasi;Ch. Bhavyasree;Ashutosh Satapathy2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) Year: 2022
- [9] Kaur, Parneet & Sikka, Karan & Divakaran, Ajay. (2017). Combining Weakly and Weakly Supervised Learning for Classifying Food Images.
- [10] Bolaños, Marc & Ferrà, Aina & Radeva, Petia. (2017). Food Ingredients Recognition Through Multi-label Learning. 10.1007/978-3-319-70742-6\_37.
- [11] Thames, Quin & Karpur, Arjun & Norris, Wade & Xia, Fangting & Panait, Liviu & Weyand, Tobias & Sim, Jack. (2021). Nutrition5k: Towards Automatic Nutritional Understanding of Generic Food. 8899-8907. 10.1109/CVPR46437.2021.00879.
- [12] WonCS(2020)Multi-scaleCNNforfine-grainedimagerecognition.IEEEAccess8:116,663–116,674.

- [13] TahirGA, LooCK(2020) An open-ended continual learning for food recognition using class incremental extreme learning machines. *IEEE Access* 8:82,328–82,346.
- [14] Liu C, Cao Y, Luo Y et al (2017) A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure. *IEEE Trans Serv Comput* 11(2):249–261.
- [15] Martini N, Foresti GL, Micheloni C (2018) Wide-slicer residual networks for food recognition. In: *IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, pp 567–576
- [16] Yanai K, Kawano Y (2015) Food image recognition using deep convolutional network with pre-training and fine-tuning. In: *IEEE International Conference on Multimedia and Expo Workshops*, IEEE, pp 1–6,
- [17] Tan M, Le Q (2019) EfficientNet: Rethinking model scaling for convolutional neural networks. In: *International Conference on Machine Learning*, PMLR, pp 6105–6114, <https://doi.org/10.48550/arXiv>.
- [18] Liang, Yanchao & Li, Jianhua. (2017). Computer vision-based food calorie estimation: dataset, method, and experiment.
- [19] Polyzotis, Neoklis et al. “Data Management Challenges in Production Machine Learning.” *Proceedings of the 2017 ACM International Conference on Management of Data* (2017)
- [20] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 265-283. Preuzetos <https://arxiv.org/abs/1603.04467>
- [21] Anandkumar, A., Ge, R., Hsu, D., Kakade, S., & Telgarsky, M. (2014). Tensor Decompositions for Learning Latent Variable Models. *Journal of Machine Learning Research*, 15(1), 2773-2832.
- [22] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- [23] Johnson, J., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1), 1-54.
- [24] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural computation*, 12(10), 2451-2471.

- [25] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.
- [26] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- [27] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
- [28] Scherer, D., Müller, A., & Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. *International Conference on Artificial Neural Networks*, 92-101.
- [29] Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359.
- [30] Howard, A. G., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., ... & Adam, H. (2019). Searching for MobileNetV3. *Proceedings of the IEEE International Conference on Computer Vision*, 1314-1324.
- [31] Zoph, B., Vasudevan, V., Shlens, J., & Le, Q. V. (2018). Learning transferable architectures for scalable image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 8697-8710.
- [32] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks?. *Advances in Neural Information Processing Systems*, 27, 3320-3328.
- [33] Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*.
- [34] Alpaydin, E. (2020). *Introduction to Machine Learning*.
- [35] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*
- [36] <https://data.nal.usda.gov/dataset/food-and-nutrient-database-dietary-studies-fnDDS>
- [37] Mureşan, H., & Oltean, M. (2018). "Fruit recognition from images using deep learning." *Acta Universitatis Sapientiae, Informatica*, 10(1), 26-42. doi:10.1515/ausi-2018-0002.
- [38] Dandekar, Mohit & Pun, Narinder & Sonbhadra, Sanjay & Agarwal, Sonali. (2020). Fruit classification using deep feature maps in the presence of deceptive similar classes.
- [39] <https://www.kaggle.com/datasets/catarinaantelo/portuguese-meals>

- [40] <https://www.kaggle.com/datasets/karkengchan/malaysia-food-11>
- [41] An image processing approach for calorie intake measurement, G.Villalobos, R.Almaghrabi, P.Pouladzadeh, 2012
- [42] <https://www.kaggle.com/datasets/moltean/fruits>
- [43] <https://www.fns.usda.gov/>
- [44] TensorFlow: Large-scale machine learning on heterogeneous distributed systems. Martìn Abadi, Ashish Agarwal, et al. 2016. <https://www.tensorflow.org/>
- [45] Chollet, F. 2015. Keras: Deep Learning library for Theano and TensorFlow. GitHub repository. <https://keras.io>
- [46] TensorFlow Datasets Documentation: <https://www.tensorflow.org/datasets>
- [47] Nair, V., & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27th International Conference on Machine Learning (ICML)*
- [48] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer. (Sigmoid function for multi-label classification)
- [49] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer. (Linear activation for regression tasks)
- [50] Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd International Conference on Machine Learning (ICML)*.
- [51] Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*.
- [52] Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. *arXiv preprint arXiv:1206.5533*
- [53] Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*.
- [54] Prechelt, L. (1998). Early Stopping — But When? In *Neural Networks: Tricks of the Trade* (pp. 55-69). Springer.
- [55] Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833).

[56] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.

## Lista slika

1. [1] <https://medium.com/@michaeldelsole/what-is-one-hot-encoding-and-how-to-do-it-f0ae272f1179>
2. [2] Treceñe, Jasten Keneth. (2019). DEVELOPMENT OF AN ANDROID BASED REAL-TIME IMAGE RECOGNITION APPLICATION USING CONVOLUTIONAL NEURAL NETWORK ALGORITHM: ADDRESSING CHALLENGES IN MOTHER TONGUE-BASED OF MULTILINGUAL EDUCATION. 10.13140/RG.2.2.18777.57446.

# IZJAVA

kojom izjavljujem s punom materijalnom i moralnom odgovornošću da sam diplomski rad s naslovom EVALUACIJA MODELA ZA ESTIMACIJU KALORIJA HRANE PRIKUPLJANJEM STANDARDIZIRANOG SKUPA PODATAKA izradio samostalno pod voditeljstvom prof. dr. sc., Saše Mladenovića, U radu sam primijenio metodologiju znanstvenoistraživačkog rada i koristio literaturu koja je navedena na kraju diplomskog rada. Tuđe spoznaje, stavove, zaključke, teorije i zakonitosti koje sam izravno ili parafrazirajući naveo/la u diplomskom radu na uobičajen, standardan način citirao sam i povezo/la s fusnotama s korištenim bibliografskim jedinicama. Rad je pisan u duhu hrvatskog jezika.

Student:



Marko Burazin