

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

Automatizacija evaluacije dizajna korisničkih sučelja

Ana Terović

Voditelj: *Igor Mekterović*

Zagreb, rujan 2023.

SADRŽAJ

1. Uvod	1
2. Dizajn korisničkog sučelja	2
2.1. Vizualna estetika i dizajn	3
2.2. Skupovi podataka	3
3. Evaluacija dizajna korisničkog sučelja	5
3.1. Metode temeljene na ocjeni korisnika	6
3.2. Algoritamske metode	7
3.2.1. Analiza vizualne složenosti temeljem metričnih modela	7
3.2.2. Heuristička analiza vizualne složenosti	8
3.2.3. Analiza vizualne složenosti na temelju strojnog učenja	8
4. Automatizacija prikupljanja informacija	10
4.1. Prikupljanje podataka iz koda	10
4.2. Prikupljanje podataka iz slike	10
5. Istaknuti radovi	11
5.1. <i>Automatsko izdvajanje i integracija metrika za web korisnička sučelja</i> .	11
5.2. <i>Evaluacija korisničkog sučelja metodama strojnog učenja</i>	13
5.3. <i>Prema poboljšanoj kreativnosti u dizajnu sučelja kroz automatiziranu procjenu upotrebljivosti</i>	14
6. Zaključak	16
7. Literatura	17
8. Sažetak	19

1. Uvod

Korisnička sučelja jedan su od najčešće korištenih sustava za interakciju. Iako je sadržaj sučelja od velike važnosti, izgled i osjećaj jednako su bitni čimbenici koji određuju kvalitetu sučelja. Važnu ulogu dizajna pokazala su i istraživanja koja govore da estetika utječe na preferencije [A. N. Tuch (2010)], percepciju upotrebljivosti [A. Sonderegger (2010)], vjerodostojnost [F. Alsudani (2009)] i performanse [C. Salimun (2010)]. Kako bi se dizajnerima i developerima olakšao proces izrade korisničkih sučelja u ovome seminaru želimo istražiti načine efikasnije procjene dizajna sučelja.

Cilj ovog seminara je istražiti mogućnosti uporabe konvencija i praksi za dobar dizajn u svrhu automatizacije evaluacije dizajna korisničkih sučelja.

Rad je podijeljen u četiri dijela. Na početku definiramo što znači dobar dizajn te kako ga možemo mjeriti. Vizualna složenost jedan je od najčešćih kriterija ocjenjivanja izgleda slike. Zato je u drugome dijelu prikazana podijela metoda analize vizualne složenosti u kontekstu korisničkih sučelja. Podijela se sastoji od metoda temeljenih na ocjeni korisnika i algoritamskih metoda, gdje su zatim algoritamske metode podijeljene na metrične, heurističke i metode utemeljene na strojnom učenju. U trećem dijelu su opisane automatizirane metode, analize koda i slike kojima možemo dobiti informacije o korisničkim sučeljima. Na kraju je napravljen sažeti pregled tri istraživačka rada koji se čine relevantni za budući rad.

2. Dizajn korisničkog sučelja

Većina formalnih definicija dobrog dizajna korisničkog sučelja i iskustva proizlaze iz radova znanstvenika Jakoba Nielsena. Jedno od istaknutijih radova je *10 heuristika upotrebljivosti za dizajn korisničkog sučelja* Nielsen (2022) koji govori da dizajn treba:

1. sadržavati vidljivost statusa sustava,
2. biti usklađen sa stvarnim svijetom,
3. omogućavati korisniku da je u kontroli i slobodan,
4. biti konzistentan i u skladu sa standardima,
5. preventirati pogreške,
6. primjenjivati pravilo prepoznavanja prije prisjećanja,
7. biti fleksibilan i efikasan za korištenje,
8. biti estetičan i minimalističan,
9. omogućiti korisnicima da prepoznaju i oporave uslijed pogrešaka,
10. imati dokumentaciju i mogućnost traženja pomoći.

Uz njegove radove, široku primjenu imaju i osnovna načela geštalt percepcije:

1. sličnost (engl. *similarity*),
2. kontinuitet (engl. *continuation*),
3. zatvorenost (engl. *closure*),
4. bliskost (engl. *proximity*),
5. zajedničko kretanje (engl. *common fate*),
6. figura i pozadina (engl. *figure/ground*).

2.1. Vizualna estetika i dizajn

Zanima nas kojim karakteristikama korisničkog sučelja možemo mjeriti dobar dizajn pa zato pogledajmo rezultate nekoliko radova koji su statistički testirali što čini dizajn korisničkog sučelja privlačnim.

Prvi rad K. Reinecke (2013) koristi vizualnu složenost (engl. *visual complexity*) i količinu boje kako bi procijenio prvi dojam estetskog izgleda web-stranice. Rezultati pokazuju da su ocjene sučelja signifikantno negativno korelirane s vizualnom složenosti, a zanimljiv podatak je da boja ima malu ulogu u prvom dojmu dizajna sučelja. U radu Aliaksei Miniukovich (2015) isto tako procjenjuju estetiku korisničkog sučelja ali ovoga puta ne samo na prvi dojam. Uz vizualnu složenost sada su i boja, rešetkasta kompozicija i količina bijelog prostora jedni od glavnih prediktora. Rad Thomas Schmidt (2018) mjeri i prvi i refleksivni dojam estetike sučelja te navode kako su oni izrazito pozitivno korelirani. Vizualna složenost se ponovno pokazala kao jaki negativno korelirani prediktor dok s druge strane, simetričnost dizajna sučelja se pokazala kao najjači pozitivno korelirani prediktor. Boja je pozitivno korelirana ako je srednjeg stupnja intenziteta.

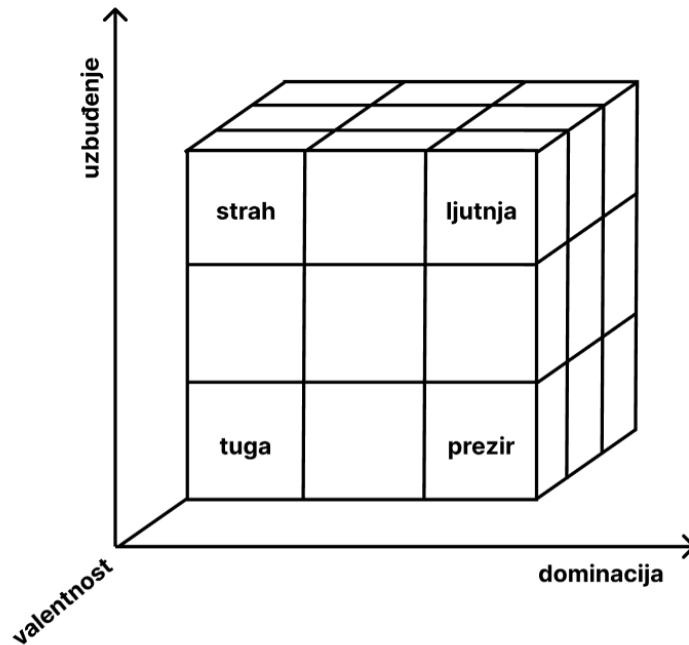
Vidimo da je vizualna složenost bitni prediktor izgleda korisničkog sučelja. Osim gore navedenih istraživanja vizualna složenost se često spominje i u drugim radovima gdje se analizira izgled slika. Zato ćemo u nastavku ovog seminara pogledati metode analize vizualne složenosti u kontekstu korisničkih sučelja kako bi napravili evaluaciju dizajna.

2.2. Skupovi podataka

Za potrebe razumijevanja ljudskih emocija prilikom percipiranja svoje okoline i gledanja slika razvijen je ogroman broj istraživanja i standardiziranih baza podataka o emocionalnim podražajima. Iako ovi skupovi podataka ne sadrže opise i slike korisničkih sučelja svejedno su korisni za razumijevanje i analizu kako ljudi percipiraju slike. Može se napraviti primjena zaključaka na temelju ovih skupova podataka na korisnička sučelja.

International Affective Picture System, IAPS je zlatni standard emocionalno i vizualno složenih slika. Često se koristi za istraživanja različitih aspekata emocionalne obrade, a sami proces uporabe IAPS skupa podataka nema utvrđenu strategiju već je prilagođen pojedinačnim istraživanjima [Jana Machajdik (2008)]. Svaka slika u bazi podataka je popraćena nizom normi (srednja vrijednost i standardna devijacija) kroz tri dimenzije: uzbuđenje (fiziološka aktivacija izazvana slikom), valentnost (ugoda i

zadovoljstvo) i dominacija (stupanj do kojeg emocionalno stanje je pod kontrolom subjekta). Ovakva trodimenzionalna struktura za mjerenje emocija, prikazana na slici 2.1, naziva se PAD model (engl. Pleasure Affect Dominance) [Mehrabian (1996)], a valjanost ovog teorijskog modela akumulirala je mnoštvo empirijskih dokaza tijekom vremena.



Slika 2.1: PAD model

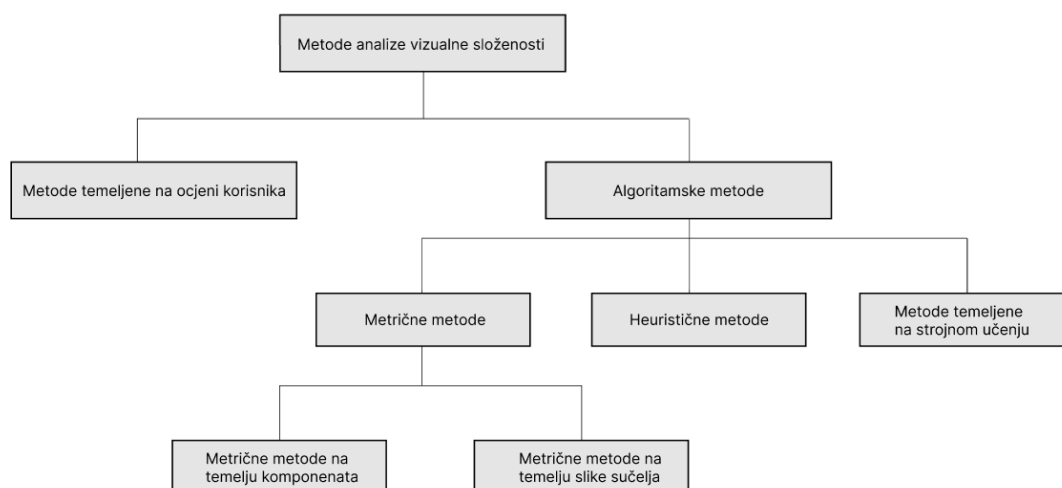
Međutim, IAPS je nastao u vrijeme prije interneta i kao takve slike podliježu ograničenjima autorskih prava koja zabranjuju njihovu upotrebu u online istraživačkim radovima. Točnije, ugovor o autorskim pravima koji prati IAPS jasno propisuje da korisnicima nije dopušteno „postaviti ih (tj. IAPS slike) na bilo koje internetske ili računalno dostupne web stranice” [Benedek Kurdi (2017)]. Open Affective Standardized Image Set, OASIS je skup podataka otvorenog pristupa koji je nastao na uzor IAPS skupu podataka s ciljem da razrješi problem ograničenosti uporabe. OASIS se temelji na dvije dimenzije: uzbuđenje i valentnost.

Razvijene su i druge baze podataka emocionalno podražajnih statičkih slika s različitim sadržajem i provjerenim normativnim ocjenama. Do danas, međunarodno najpriznatije baze podataka su Geneva Affective Picture Database, GAPED [Elise S. Dan-Glauser (2011)] i Nencki Affective Picture System, NAPS [Artur Marchewka (2014)]. GAPED slike su ocjenjivane prema valentnosti, uzbuđenosti i usklađenosti prikazane scene s unutarnjim (moralnim) i vanjskim (pravnim) normama. NAPS ocjenjuje slike kroz tri dimenzije: valentnost, uzbuđenje i pristup-izbjegavanje korištenjem kompjuteriziranih bipolarnih semantičkih kliznih ljestvica.

3. Evaluacija dizajna korisničkog sučelja

Evaluacija dizajna korisničkog sučelja biti će napravljena metodama analize vizualne složenosti. Vizualna složenost kao koncept mjerila izgleda slike ima mnogo interpretacija, a s time i mnogo primjena. Teško je donijeti jasnu definiciju budući da je pojam vizualne složenosti povezan s ljudskom percepcijom ali bitno je razumijeti koje komponente korisničkog sučelja utječu da ga čovjek percipira složenim.

Predstaviti ćemo podjelu metoda analize vizualne složenosti u kontekstu korisničkih sučelja prikazanu na slici 5.1 po uzoru na rad Eren Akça (2021). Rad je nakon identificiranja i analize 85 istraživačkih radova, napravio pregled i grupirao metode analize vizualne složenosti te je u skladu s tim prikazana taksonomija.



Slika 3.1: Podjela metoda analize vizualne složenosti

3.1. Metode temeljene na ocjeni korisnika

Metode temeljene na ocjeni korisnika se oslanjaju na ljudskoj procjeni vizualnih objekata. Korisniku je predstavljena slika nakon čega kroz upitnik ili razgovor iznosi svoja mišljenja. Pozitivna strana ovakvog pristupa je što je instiktivan, ne zahtjeva matematičke izračune te rezultati mogu nadilaziti očekivane ishode. S druge strane, ovakve metode imaju subjektivne rezultate jer se baziraju na ljudskim procjenama, vremenski su neefikasne i nepraktične u usporedbi s algoritamskim metodama.

Prvi korak je utvrditi hipotezu koja će se testirati. Kao što je prethodno rečeno, postoje mnoge interpretacije vizualne složenosti te s obzirom na nju će se definirati i hipoteza istraživanja. Nakon određene hipoteze gradi se upitnik. Upitnik treba biti jasan, pažljivo formuliran i mjerljiv kako bi rezultati istraživanja bili što bolji. Kvalitetu istraživanja isto tako možemo poboljšati odabirom raznolike grupe korisnika koji će sudjelovati u istraživanju. Na kraju se podaci dobiveni iz istraživanja statistički obrađuju i analiziraju kako bi se izvuklo korisno znanje.

Rad Eleni Michailidou (2008) predstavlja evaluaciju koja ispituje kako korisnici povezuju estetski izgled web stranice s njezinom vizualnom složenošću. Istraživanje je otkrila jaku povezanost između vizualne složenosti, estetike i izgleda web stranice. Izbornici, slike, riječi, linkovi i gornji lijevi kut (engl. top-left corner, TLC) definirani su kao ključne varijable strukture web-stranice. Određene su tri hipoteze:

1. Broj izbornika, slika, vidljivih poveznica, riječi i TLC-ova koje web-stranica ima, pozitivno je povezan s razinom vizualne složenosti stranice.
2. Korisnička percepcija estetskih karakteristika web stranice s obzirom na organizaciju, preglednost, čistoću, zanimljivost i ljepotu linearno je povezana s brojem izbornika, slika, vidljivih poveznica, riječi i TLC-ova koje web-stranica ima.
3. Korisnička percepcija vizualne složenosti web-stranice povezana je s estetskim kvalitetama koje ona predstavlja. Odnosno, što je web stranica organiziranija, jasnija, čišća, zanimljivija i ljepša, korisnik je percipira manje vizualno složenom.

Prva hipoteza djelomično je potvrđena odnosno eksperiment je otkrio pozitivan i značajan odnos između vizualne složenosti stranice i broja njezinih slika, vidljivih poveznica, riječi i TLC-ova, ali ne i s brojem izbornika. Druga hipoteza je podržana za sve strukturne elemente osim za broj izbornika. Za treću hipotezu eksperiment je otkrio da samo zanimljivost nije povezana s razinom složenosti. Odnosno, analiza podataka pokazala je da što je stranica uređenija, jasnija, čišća i ljepša, to su je sudionici vizualno jednostavnije percipirali.

Neka od zanimljivih observacija istraživanja su da su najorganiziranije, najjasnije i najljepše stranice bile među prve tri vizualno jednostavnije stranice. Veliku ocjenu zanimljivosti dobila je jedna od vizualno najsloženijih stranica.

3.2. Algoritamske metode

Algoritamske metode koriste konkretne mjere izračuna vizualne složenosti. Koristeći formalne operacije u preddefiniranim izračunima ovakve metode su objektivne, konzistentne i determinističke. Uz navedene, bitni pozitivni faktor algoritamskih metoda je vremenska efikasnost gdje nakon određivanja modela računanje rezultata zahtjeva minimalno vremena. Negativna strana korištenja algoritamskih metoda je da nereprezentiraju mentalnu i kognitivnu percepciju korisnika kao što to čine metode temeljene na ocjeni korisnika. Algoritamske metode se dijele na metrične, heurističke i metode temeljene na strojnom učenju.

3.2.1. Analiza vizualne složenosti temeljem metričnih modela

Metrični modeli utemeljeni su na točno zadanim matematičkim izrazima koji definiraju apstraktne koncepte vizualne složenosti. Iz ovoga proizlazi i glavna motivacija izrade metričnih modela, a to je da čine apstraktne koncepte mjerljivima, dokazivima te nam daju mogućnost da ih verificiramo. Napravljena je distinkcija dviju kategorija metričnih modela: metrični modeli temeljeni na komponentama sučelja i metrični modeli temeljeni na slici sučelja. Razlika je u tome da modeli temeljeni na komponentama baziraju svoje matematičke izraze na karakteristikama komponentama, na primjer gledati će se broj ili veličina neke komponente. S druge strane, modeli temeljeni na slici sučelja će svoje matematičke izraze definirati karakteristikama slike sučelja, na primjer stupanj kompresije ili veličina slike koja prikazuje korisničko sučelje.

Iz prethodne klasifikacije modela možemo vidjeti da matematički izrazi moraju biti determinirani s obzirom na karakteristike korisničkih sučelja. Kako bi ostvarili dobre rezultate potrebno je prethodno imati dobro razumijevanje koji elementi sučelja imaju veliki utjecaj na ljudsku percepciju. U prvome koraku ćemo prethodno određene karakteristike sučelja utjeloviti i pretvoriti u matematičke izraze. Primjer toga može biti uzimanje srednje vrijednosti i standardne devijacije veličine komponentata na sučelju. U sljedećem koraku se procesiranjem slike radi izdvajanje obilježja (engl. *feature extraction*) nakon čega matematičkim izrazima radimo izračun vizualne složenosti. Zadnjim korakom radimo verifikaciju rezultata kvalitativnim evaluacijama.

Cilj rada Aliaksei Miniukovich (2014) je smanjiti potreban trud dizajnera da izgrade upotrebljiva i vizualno privlačna korisnička sučelja sistematičnim pristupom korište-

njem objektivnih kriterija, tj. metrika. Istraženo je pet karakteristika: (1) vizualni šum, (2) simetrija, (3) gustoća kontura, (4) kontrast figure i tla i (5) varijabilnost boja. Sukladno njima napravljeno je šest metrika, a kao primjer formula 3.1 predstavlja matematički izraz kojim su računali simetriju. Rezultati istraživanja su objasnili 51% ocjena estetike i 50% ocjena složenosti.

$$sym_{norm} = \frac{k_{sym}}{k_{all}} * \left(\frac{(k_{all} - 1) * s_a}{s_s} \right)^{-1} \quad (3.1)$$

3.2.2. Heuristička analiza vizualne složenosti

Heuristički algoritmi određuju kvalitetu dizajna korisničkih sučelja i mjere vizualnu složenost na temelju generiranih pravila procjene. Ova pravila predstavljaju karakteristike kvalitativnih podataka uključujući ocjene korisnika, korisničke profile i uočene nedostatke. Heurističke metode su korisne kada problem ne možemo riješiti determinističkim metodama odnosno iako ne možemo pronaći točno rješenje heurističkim metodama možemo pronaći optimalno rješenje.

Nakon prikupljanja podataka, generira se skup slučajno inicijaliziranih pravila. Zatim započinje iterativni heuristički proces. Sve dok se ne dostigne zadovoljavajući uvjet, u svakoj iteraciji evolucija jednog skupa pravila se provjerava funkcijom dobrote. Nakon izvršenja svih iteracija, dobiveni skup pravila smatra se najboljim pravilima procjene i koristi se za mjerenje kvalitete korisničkih sučelja.

Rad Gasmi Ines (2017) koristi genetički algoritam kako bi stvorio optimalni skup pravila za procjenu kvalitete mobilnih sučelja. Postignut je rezultat s točnosti od 70%.

3.2.3. Analiza vizualne složenosti na temelju strojnog učenja

Metode strojnog učenja grade svoje modele treniranjem na skupovima podataka. Kako bi se postigla dobra generalizirana rješenja skupovi podataka trebaju biti što veći. Uobičajena ograničenja postojećih metoda su nedostatak skupova podataka za korisnička sučelja te manjak istraživanja koja koriste metode strojnog učenja za analizu korisničkih sučelja, osobito s novijim dubokim neuronskim mrežama. Dobra strana je da nakon što se model nauči, nove predikcije se dobivaju izrazito brzo, a ako je model dobro naučen točnost će biti visoka.

Nakon skupljenog velikog skupa podataka, potrebno je podatke označiti te ih podijeliti na skupove za testiranje, treniranje i validaciju. U sljedećem koraku se biraju značajke na temelju kojih će model raditi predikcije. Model zatim treniramo na skupu za treniranje dok ne postignemo zadovoljavajuće rezultate, a na kraju se ti rezultati potvrđuju primjenom prethodno natreniranog modela na skupu za testiranje.

U svrhu procjene vizualne složenosti korisničkog sučelja rad Ou Wu (2013) koristi strojno učenje. Za pripreme podataka napravljeno je izdvajanje značajki (engl. *feature extraction*) te podijela na skup za treniranje i testiranje. Tijekom izdvajanja značajki napravljena je strukturalna i vizualna analiza rezultirajući sa skupom podataka s 44 značajke. Značajke su opisivile boje, raspored, slike i linkove. Isprobani su modeli RF (engl. *Random Forest*) , SVM (engl. *Support Vector Machine*) i neuronske mreže. Najbolje rezultate dao je SVM-model s vrijednosti $R^2 = 0.7964$. Analizom značajki utvrđeno je da su neke od najbitnijih značajki: veličina slike, broj slika, prosječna svjetlina, varijanca svjetline, širina i visina stranice te veličina fonta.

4. Automatizacija prikupljanja informacija

Analiza dizajna korisničkog sučelja zahtjeva prikupljanje podataka koji govore o izgledu sučelja. Kako bi prikupili te podatke možemo koristiti metode analize koda i slike. Ove metode mogu imati i primjenu u izgradnji skupova podataka na temelju kojih možemo trenirati modele. Napravimo sada kratki pregled navedenih metoda.

4.1. Prikupljanje podataka iz koda

Analiza izvornog koda je automatizirano testiranje izvornog koda. U osnovi se radi o otklanjanje pogrešaka koda gdje je cilj je pronaći pogreške koje programeru možda nisu očite. Analizatori koda rade koristeći pravila koja mu govore što treba tražiti [Techopedia (2022)]. Primjeri alata za evaluaciju korisničkih sučelja koji koriste analizu koda su Wave [Wave (2022)] i Axe [Axe (2022)]. Oba alata analiziraju CSS i JS kod web-stranica kako bi uočili nedostatke. Alati će na primjer uočiti da slici nedostaje naslov ili da je font postavljen na premalu veličinu.

4.2. Prikupljanje podataka iz slike

Izvlačenje značajki (engl. *feature extraction*) je metoda koja se često spominje u radovima o evaluaciji korisničkih sučelja. To je važan korak u izgradnji bilo koje klasifikacije uzoraka i cilj joj je izdvajanje relevantnih informacija koje karakteriziraju svaku klasu. U ovom procesu relevantne značajke izdvajaju se iz objekata kako bi se formirali vektori značajki. Te vektore značajki zatim koriste klasifikatori za prepoznavanje i povezivanje ulazne varijable s ciljnom izlaznom varijablom. Izvlačenje značajki je proces dohvaćanja najvažnijih podataka iz neobrađenih podataka [Gaurav Kumar (2014)]. U kontekstu ovoga seminara, ovom metodom možemo na primjer odrediti koliko i kakve sve elemente imamo na korisničkom sučelju. Postoji veliki broj implementiranih rješenja koja su slobodna za uporabu.

5. Istaknuti radovi

Predstaviti ćemo pobliže tri rada koji se čine od velikog interesa za kasniji nastavak rada u ovome području. U zaključku rada Eren Akça (2021) izneseno je da svaka kategorija metoda analize vizualne kompleksnosti ima svoje mane, no generalizirana rješenja bazirana na strojnom učenju se čine kao novi obećavajući smjer istraživanja za razvoj učinkovitih pristupa. Vođeni ovom izjavom i činjenicom da je to područje još uvijek poprilično neistraženo sva tri rada koristiti će metode strojnog učenja za analizu korisničkog sučelja. Radovi su posloženi prema kronološkom poretku publikacije.

5.1. *Automatsko izdvajanje i integracija metrika za web korisnička sučelja*

Automatsko izdvajanje i integracija metrika za web korisnička sučelja (engl. *Auto-Extraction and Integration of Metrics for Web User Interfaces*) Maxim Bakaev (2018) je rad koji je nastao kada su autori primjetili da u području procjene kvalitete korisničkih sučelja započinje prijelaz s analize koda (HTML/CSS) na pronalaženje vizualnih atributa web-stranica na temelju tehnika prepoznavanja slika. Motivaciju za istraživanje ovog područja vide jer smatraju da kod koji nema mana ne mora nužno rezultirati visokom kvalitetom uporabe. Tehnike analize slike jamče da analizator radi s istim sučeljem koje korisnik doživljava, dok se HTML/CSS kod može različito prikazati u različitim preglednicima i okruženjima.

Podatke o korisničkom sučelju dobivaju na oba načina, analizom koda i analizom slike. Analizator uzima snimku zaslona korisničkog sučelja kao ulaz i pokušava identifikirati elemente od kojih je sučelje formirano. Ako je korisničko sučelje web korisničko sučelje, može se napraviti dodatna DOM analiza na temelju HTML/CSS izvornog koda. Kako bi se dobili bolji rezultati, slike su pretvorene u crno-bijelo te je napravljena dodatna analiza i detekcija pravokutnih površina, teksta, posebnih elemenata poput padajućih izbornika, te analiza kompozitnih struktura.

Dobiveno je sveukupno 7 metrika za opis korisničkog sučelja:

1. broj svih identificiranih elemenata, $VA_{Elements}$
2. broj različitih elemenata, VA_{Vocab}
3. stopa kompresije, $VA_{Compress}$
4. index vizualne kompleksnosti, VA_{Ivc} (5.1)
5. površina svih tekstualnih elemenata, VA_{Text}
6. površina svih grafičkih elemenata, VA_{Other}
7. bijela površina, VA_{White}

$$VA_{Ivc} = VA_{Elements} * \log_2(VA_{Vocab}) * VA_{Compress} \quad (5.1)$$

Napravljena je linearna regresija kako bi se odredila vizualna kompleksnost korisničkog sučelja. Rezultati pokazuju da model objašnjava 64.7% varijance te da je prilagođeni koeficijent determinancije (engl. *adjusted R squared*) $R^2 = 0.552$. Iako R^2 ne daje izrazito dobar rezultat, autori argumentiraju kako glavni cilj rada nije bio izgraditi točne prediktivne modele, već usporediti one izrađene s različitim skupovima metrika. U sklopu rada napravljena je i detaljna statistička analiza koja je pokazala da dobivene mjere analizatora za prethodno navedene metrike nisu bile poprilično ispravne. Točnije, analizator je krivo identificirao različite vrste elemenata, VA_{Vocab} te je krivo određivao površinu ispod elemenata, VA_{Other} .

Poboljšanje rada vidimo kroz poboljšanje analizatora. Rad je izašao 2018. godine što daje mogućnost da su se alati i metode za bolju detekciju elemenata i analizu slike znatno poboljšali. Isto tako bi se mogli isprobati neki drugi modeli strojnog ili dubokog učenja.

5.2. *Evaluacija korisničkog sučelja metodama strojnog učenja*

Rad *Evaluacija korisničkog sučelja metodama strojnog učenja* (engl. *User Interface Evaluation with Machine Learning Methods*) Mao (2019) je motiviran iz dva razloga: (1) učestalost korištenja korisničkih sučelja u svakodnevnom životu te (2) dugotrajnost procesa izrade korisničkih sučelja. Za ocjenu korisničkog sučelja se koriste dva aspekta, upotrebljivost i korisnost s time da je upotrebljivost proširena na još dva aspekta, efikasnost i zadovoljstvo. Autor detaljnije argumentira potrebu korištenja metoda strojnog učenja. Fokus u radu nije bio otkriti heuristiku dizajna korisničkog sučelja matematičkim izrazima, već pronaći način kako koristiti podatke za pronalaženje obrazaca potencijalnih ili nepoznatih heuristika za evaluaciju korisničkog sučelja. Stoga bi potrebna metoda implementacije trebala imati sposobnost otkrivanja implicitnih obrazaca iz podataka, što je upravo specijalnost obitelji metoda strojnog učenja.

Skup podataka sadrži podatke o lokaciji i dimenzijama svake komponente korisničkog sučelja. Implementirana su dva modela. Prvi model koristi SVM te se trenira kako bi pronašao obrasce dizajna korisničkih sučelja. Drugi model je treniran kako bi oponašao interakciju čovjeka sa sučeljima za što se koristi duboki model podržanog učenja engleskog naziva deep Q learning.

Prema rezultatima za prvi model točnost klasifikacije bila je preko 80%. U sklopu analize iznesena su tri zanimljiva zaključka o dizajnu korisničkih sučelja:

1. Sve dok korisničko sučelje nema očitih problema s upotrebljivošću korisnici ne mogu osjetiti razliku malih promjena, tj. treba izbjegavati očite probleme upotrebljivosti.
2. Aplikacije (engl. *widget*) postavljeni bliže gornjem lijevom kutu imaju višu razinu zadovoljstva.
3. S obzirom na zadovoljstvo, značajke su ocijenjene od više do niže ocijene ako su okomitih dimenzija, horizontalnih dimenzija, horizontalnog položaja i okomitog položaja.

Rad ima prostora za poboljšanje proširenjem skupa značajki. Korišteni su samo podaci o lokaciji i dimenzijama komponenata no po uzoru na druge radove mogle bi se koristiti i druge značajke kao što su visina i širina cijelog područja u kojem se nalazi korisničko sučelje, boja pozadine i svake komponente korisničkog sučelja, informacije o tipu korisničkog sučelja (desktop ili mobilno korisničko sučelje) i svake komponente (gumb, tekst, okvir itd.).

5.3. *Prema poboljšanoj kreativnosti u dizajnu sučelja kroz automatiziranu procjenu upotrebljivosti*

Prema poboljšanoj kreativnosti u dizajnu sučelja kroz automatiziranu procjenu upotrebljivosti, (engl. *Towards Enhanced Creativity in Interface Design through Automated Usability Evaluation*) Snehal Dhengre (2020) noviji je rad u području automatizacije evaluacije korisničkih sučelja. Cilj rada je istražiti i ocijeniti upotrebu strojnog učenja za predviđanje mjera upotrebljivosti sučelja mobilnih aplikacija. Korištene su konvolucijske neuronske mreže (engl. *convolutional neural network*, CNN) za predviđanje tri mjere upotrebljivosti: pravilnost (engl. *regularity*), složenost (engl. *complexity*) i dodirljivost (engl. *touchability*).

Korišten je RICO skup podataka [of Illinois] koji sadrži vizualne, tekstualne, strukturalne i interakcijske opise dizajna korisničkih sučelja mobilnih aplikacija. S obzirom na vrijednosti iz skupa podataka definirane su matematičke formule koje opisuju pravilnost, složenost i dodirljivost. Konvolucijska mreža se sastoji od dvije zajedničke konvolucije nakon čega se za svaku mjeru koristi posebna konvolucijska mreža s potpuno povezanim slojem na kraju.

Na testnom skupu su postignuti sljedeći rezultati za R^2 : pravilnost 88.57%, složenost 89.11%, dodirljivost 89.18%. Ovo su impresivni rezultati ako gledamo prethodno spomenuta istraživanja.

Ideja za proširenje ovog istraživanja je generalizirati ga na web-sučelja. Problem ove ideje je u tome što ne postoji skup podataka koji sadrži opise dizajna web korisničkih sučelja. Nedostatak adekvatnog skupa podataka vjerojatno je jedan od glavnih razloga slabog razvoja područja automatizacije evaluacije korisničkih sučelja.

	Vizualna složenost	Boje	Kompozicija	Bijeli prostor	Slike	Poveznice	Riječi	TLC	Raznolikost elemenata	Kompresija
Reinecke (2013)	+									
Aliaksei Miniukovich (2015)	+	+	+	+						
Thomas Schmidt (2018)	+		+							
Eleni Michailidou (2008)					+	+	+	+		
Aliaksei Miniukovich (2014)		+	+	+						
Ou Wu (2013)		+	+		+	+				
Maxim Bakaev (2018)	+			+	+		+		+	+
Mao (2019)			+						+	

Slika 5.1: Značajke pojedinih radova

6. Zaključak

Korisnička sučelja nisu samo most između korisnika i aplikacije nego su i način na koji aplikacija može privući korisnika svojim izgledom. Izrada dizajna korisničkih sučelja dugotrajan je i težak proces, a ovim seminarom želimo istražiti načine na koji ga možemo skratiti i olakšati.

Cilj seminara je iskoristiti znanstveno priznate definicije dobrog dizajna i iskustva uporabe korisničkog sučelja kako bi automatizirali njihovu evaluaciju.

Na početku rada je predstavljeno što čini dobar dizajn kroz definicije i druge znanstvene radove. U drugome dijelu opisane su metode kojima kvantificiramo dobar dizajn korisničkog sučelja. Ukratko su opisane metode analize koda i slike kako bi se automatiziralo prikupljanje podataka o korisničkim sučeljima. Na kraju su istaknuta tri rada koji bi mogli biti uzor za daljnje istraživanje ovoga područja.

Metoda analize vizualne složenosti na temelju strojnog učenja ističe se kao područje s puno potencijala ali malo napravljenog istraživanja. Rad Snehal Dhengre (2020) koji koristi duboko učenje pokazuje izvrsne rezultate za procjenu upotrebljivosti mobilnih sučelja. Problem proširenja takvog modela na web korisnička sučelja je nedostatak skupova podataka koji sadrže opise izgleda web-stranica. Potencijalno rješenje ovog problema možemo vidjeti u radovi Maxim Bakaev (2018) i Mao (2019) koji ne koriste postojeće skupove podataka već prikupljaju informacije analizom koda i slike.

7. Literatura

- K. Opwis A. N. Tuch, J. a. Bargas-Avila. Symmetry and aesthetics in website design: It's a man's business. *Comput. Human Behav.*, 2010.
- J. Sauer A. Sonderegger. The influence of design aesthetics in usability testing: effects on user performance and perceived usability. *Appl. Ergon.*, 2010.
- Antonella De Angeli Aliaksei Miniukovich. Quantification of interface visual complexity. *ResearchGate*, 2014.
- Antonella De Angeli Aliaksei Miniukovich. Computation of interface aesthetics. *Conference on Human Factors in Computing Systems*, 2015.
- Katarzyna Jednoróg Artur Marchewka, Łukasz Żurawski. The nencki affective picture system (naps): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, 46:596—610, 2014.
- Axe. deque. axe™ - the standard in accessibility testing,. <https://www.deque.com/axe/>, 2022.
- Mahzarin R. Banaji Benedek Kurdi, Shayn Lozano. Introducing the open affective standardized image set (oasis). *Behavior Research Methods*, 49:457—470, 2017.
- D. R. Simmons C. Salimun, H. C. Purchase. The effect of aesthetically pleasing composition on visual search performance. *Conf. Human-Computer Interact.*, 2010.
- Sean Bechhofer Eleni Michailidou, Simon Harper. Visual complexity and aesthetic perception of web pages. *ResearchGate*, 2008.
- Klaus R. Scherer Elise S. Dan-Glauser. The geneva affective picture database (gaped): a new 730-picture database focusing on valence and normative significance. *Behavior Research Methods*, 43, 2011.
- Omer Ozgür Tanriover Eren Akça. A comprehensive appraisal of perceptual visual complexity analysis methods in gui design. *Science Direct*, 2021.

- M. Casey F. Alsudani. The effect of aesthetics on web credibility. *British Computer Society*, 2009.
- Chouchane Mabrouka Gasmi Ines, Soui Makram. Evaluation of mobile interfaces as an optimization problem. *Procedia Computer Science*, 112:235–248, 2017.
- Pradeep Kumar Bhatia Gaurav Kumar. A detailed review of feature extraction in image processing systems. *iEEE*, 2014.
- Allan Hanbury Jana Machajdik. Affective image classification using features inspired by psychology and art theory. <https://www.imageemotion.org>, 2008.
- et al. K. Reinecke. Predicting users first impressions of website aesthetics with quantification of perceived visual complexity and colorfulness. *Conference on Human Factors in Computing Systems*, 2013.
- Yanxun Mao. User interface evaluation with machine learning methods. 2019.
- Vladimir Khvorostov Maxim Bakaev, Sebastian Heil. Auto-extraction and integration of metrics for web user interfaces. *Journal of Web Engineering*, 17:561–590, 2018.
- Albert Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14:261–292, 1996.
- Jakob Nielsen. *10 usability heuristics for user interface design*, 2022. URL <https://www.nngroup.com/articles/ten-usability-heuristics/>.
- The University of Illinois. Rico dataset. <https://interactionmining.org/rico>.
- Lei Shi Ou Wu, Wwiming Hu. Measuring the visual complexities of web pages. *Acm Transactions on Web*, 2013.
- Farzaneh Oghazian Snehal Dhengre, Jayant Mathur. Towards enhanced creativity in interface design through automated usability evaluation. *Eleventh International Conference on Computational Creativity*, 2020.
- Techopedia. *Techopedia*, 2022. URL <https://www.techopedia.com/definition/29997/source-code-analysis>.
- Christian Wolf Thomas Schmidt. The influence of user interface attributes on aesthetics. *i-Com*, 2018.
- Wave. *Wave*, 2022. URL <https://wave.webaim.org/api/>.

8. Sažetak

Korisnička sučelja integralan su dio izgradnje web-aplikacija. Korisnici koji se ne bave programiranjem najčešće neće razumjeti kompleksnost programa iza sučelja, ali zato će im osim funkcionalnosti biti bitan i njen izgled, preglednost i moći će cijeniti njen dizajn. Za dizajn korisničkih sučelja postoje predložene konvencije i prakse koje imaju cilj doprinijeti vizualnome izgledu i iskustvu uporabe sučelja. Kroz ovaj seminar istražiti ćemo mogućnost uporabe konvencija i praksi za dobar dizajn u svrhu automatizacije evaluacije dizajna sučelja. Seminar će sadržavati uvod u područje te upoznavanje s već predloženim rješenjima evaluacije i automatizacije testiranja korisničkih sučelja. Neki od pristupa primjenjuju metode strojnog učenja, analizu koda, vizualnu segmentaciju slike, i dr.