

Personalizirani modeli za napovedovanje poslabšanja kroničnega srčnega popuščanja

Personalized Models to Detect Worsening of Condition of Chronic Heart Failure Patients

Martin Gjoreski, Anton Gradišek, Borut Budna
Odsek za inteligentne sisteme, Institut Jožef Stefan
martin.gjoreski@ijs.si

Gregor Pogljajen
Univerzitetni klinični center Ljubljana
gregor.pogljajen@ukclj.si

POVZETEK

Kronično srčno popuščanje, pri katerem srce ne more črpati dovolj krvi, da bi zadostilo potrebam telesa, spada med kronične bolezni, ki pogosto prizadenejo predvsem starejše od 65 let. Ob poslabšanju stanja je pogosto potrebna hospitalizacija, kar predstavlja breme tako za pacienta kot za zdravstveni sistem. Srčno popuščanje se do neke mere lahko zazna na podlagi spremembe srčnih tonov, kar smo že pokazali v predhodnih študijah, kjer smo razvili metodo, ki na podlagi zvočnih posnetkov ločuje med zdravimi posamezniki in pacienti z epizodo poslabšanja srčnega popuščanja. Naš cilj je razviti telemedicinsko metodo, ki na podlagi analize srčnih tonov z metodami strojnega učenja prepozna poslabšanje. V prispevku predstavimo začetne rezultate gradnje personaliziranega modela, kjer primerjamo posameznike v dveh fazah srčnega popuščanja – v rekompenzirani in dekompenzirani fazi. Eksperiment smo izvedli na posnetkih 22 posameznikov in dosegli 91 % natančnost pri ločevanju med obema fazama na podlagi ene značilke.

Ključne besede

Srčno popuščanje, analiza zvoka, strojno učenje.

ABSTRACT

Chronic heart failure, where the heart cannot pump enough blood to supply the body's needs, is a chronic disease that affects especially people above 65. Worsening of the condition often requires hospital treatment, which is a burden both for the patient and for the health system. To some degree, it is possible to detect heart failure based on the changes of heart sounds, which we demonstrated in a previous study. We developed a method which, based on sound recordings, distinguishes between patients and healthy individuals. Our aim is to develop a telemonitoring method that will use advanced machine-learning methods on sound recordings to detect onset of worsening. Here we present the initial results of building a personalized model, where we compare individuals in two stages of heart failure – in decompensated and recompensated phase. The experiment was carried out using recordings of 22 patients. Based on a single feature, we obtained the accuracy of 91%.

Keywords

Chronic heart failure, sound analysis, machine learning.

1. UVOD

Kronično srčno popuščanje (SP, angleško Chronic Heart Failure, CHF) je ena od epidemij sodobnega časa. Gre za kronično

progresivno bolezen, kjer srce ni sposobno črpati dovolj krvi, da bi zadostilo potrebam telesa, kar se pri pacientih izraža kot izrazitejše poslabšanje telesne zmogljivosti, težje dihanje ter otekanje v okončine in/ali trebuh. Bolezen v razvitem svetu prizadene 1-2 % splošne populacije ter do 10 % populacije, starejše od 65 let. Po ocenah trenutno (podatek iz leta 2015) živi na svetu več kot 23 milijonov bolnikov s SP, od tega več kot 5 milijonov v ZDA. Študije ocenjujejo, da tam strošek za skrb za bolnike znaša 31 milijard dolarjev, kar bo do leta 2030 naraslo na 70 milijard [1][2].

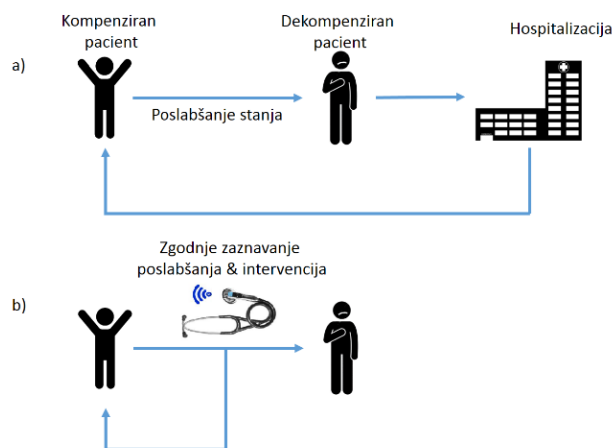
V kliničnem poteku SP se tipično izmenjujejo obdobja dobrega počutja (tj. kompenzirane faze) in obdobja poslabšanja SP (tj. faze dekompenzacije), ko pacient opaža zgoraj opisane simptome. Ob poslabšanju SP je pogosto potrebna hospitalizacija, kjer se bolnikom pomaga z intravenoznimi zdravili. Če se poslabšanje SP zazna dovolj zgodaj, lahko z ustreznimi terapevtskimi ukrepi, kot so omejitve vnosa tekočine in povečanjem odmerka diuretikov, bolniku prihranimo hospitalizacijo.

Izkušen zdravnik lahko poslabšanje SP prepozna ob pregledu bolnika, pogosto se poslabšanje izraža kot sprememba oz. pojavljanje dodatnih srčnih tonov (tipično tretji ton, ki se pojavi 0,1-0,2 s za drugim tonom), kar zdravnik spremlja s stetoskopom (fonokardiografija). Raziskave v zadnjem času kažejo, da se nekateri fiziološki parametri (npr. dodatni srčni toni, porast tlakov v pljučnem krvotoku) pričnejo spreminjati že nekaj tednov preden pride do klinično očitnega poslabšanja SP. Prve analize [3] že kažejo, da zgodnje ugotavljanje spreminjanja teh parametrov lahko pomeni precej učinkovitejšo obravnavo SP, saj je epizod polno razvitega poslabšanja srčnega popuščanja manj, s tem pa tudi manj potrebe po hospitalni obravnavi.

V pričujočem prispevku raziskujemo možnosti zaznavanja poslabšanja SP na podlagi analize srčnih tonov z metodami umetne inteligence. Raziskava je del daljšega projekta, cilj katerega je razviti metodo ali telemedicinsko napravo, s katero bo uporabnik lahko sam spremljal svoje zdravstveno stanje in ob nakazanem poslabšanju SP kontaktiral zdravnika, kar shematsko prikazuje Slika 1.

V preteklih analizah [4][5] smo razvili metodo, s katero uspešno ločujemo med posnetki srčnih tonov zdravih posameznikov ter bolnikov v dekompenzirani fazi SP. Metoda je sestavljena iz filtriranja, segmentacije, izračuna značilk in gradnje klasifikacijskih modelov strojnega učenja. Klasifikatorje smo kombinirali z uporabo pristopa zlaganja (stacking), kjer so osnovni klasifikatorji klasificirali posamezne segmente, meta-klasifikator pa je klasificiral celotne posnetke, saj vemo, da vsi segmenti nekega posnetka pripadajo istemu razredu. Metoda je v primerjavi z večinskimi klasifikatorjem dosegla 15 % višjo natančnost.

V pričujočem prispevku se ukvarjamo z vprašanjem personalizacije. Zanima nas, če lahko ločujemo med posnetki pri posameznikih v različnih fazah SP. Na tej stopnji primerjamo posnetke pacientov v dekompenzirani fazi, ob sprejetju v bolnišnico, ter v rekompenzirani fazi, ko je pacient odpuščen iz bolnišnice.



Slika 1. Shematski prikaz razvoja poslabšanja srčnega popuščanja. V zgornjem primeru mora pacient zaradi poslabšanja SP v bolnišnico (hospitalizacija), v spodnjem primeru poslabšanje zaznamo dovolj zgodaj, da lahko z ustrežno intervencijo preprečimo poslabšanje in se hospitalizaciji izognemo.

2. MATERIALI IN METODE

2.1 Baza podatkov

V študiji smo obravnavali posnetke 22 bolnikov s kroničnim srčnim popuščanjem, ki so bili hospitalizirani zaradi poslabšanja zdravstvenega stanja. Prvi posnetek je nastal v dekompenzirani fazi (ob hospitalizaciji), drugi pa v rekompenzirani fazi, ob odpustu iz bolnišnice.

Za snemanje smo uporabili profesionalni medicinski digitalni stetoskop 3M Littmann Electronic Stethoscope Model 3200. Posnetek, fonokardiogram, je bil vedno sneman na Erbovi točki, nad tretjim medrebrnim prostorom levo od grodnice. Posamezen posnetek je dolg do 30 s, kar je omejeitev stetoskopa.

Pred začetkom študije smo pridobili pozitivno mnenje Komisije za medicinsko etiko.

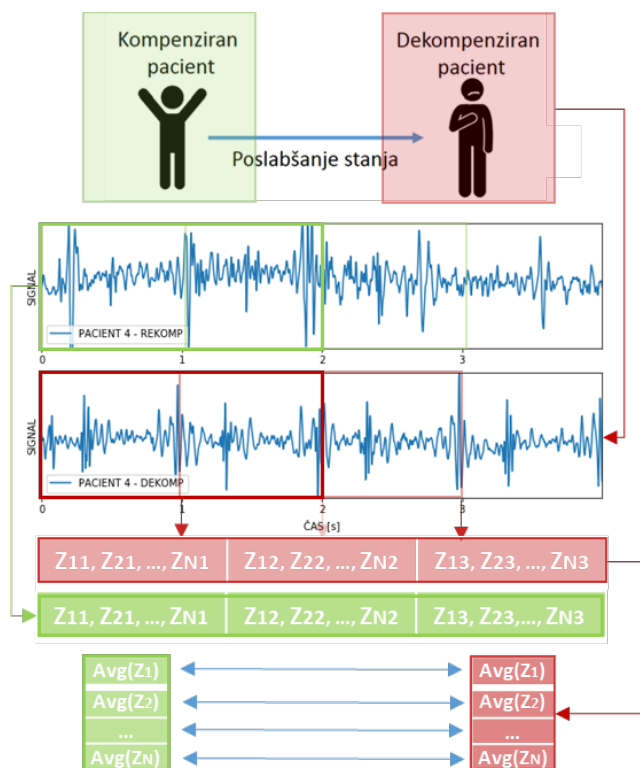
2.2 Metoda

Za analizo smo uporabili orodje za klasifikacijo zvokov [6], ki smo ga razvili na platformi za delo s podatki Orange. Orodje za ekstrakcijo značilk uporablja programsko opremo OpenSMILE [7], ki je bila prvotno razvita leta 2009 za prepoznavanje čustev na podlagi zvoka, vendar se je kasneje začela uporabljati tudi v bolj splošne namene, kot je pridobivanje informacij iz glasbenih posnetkov.

Shema metode je predstavljena na Sliki 2. Za vsakega pacienta smo imeli 2 posnetka, enega v rekompenzirani fazi (zelena) in drugega v dekompenzirani fazi (rdeča). Najprej smo vse posnetke segmentirali z uporabo tehnike drsečega okna. Na Sliki 2 imamo posnetek dolžine štirih sekund. Za velikost drsečega okna dve sekundi in 50 % prekrivanja (dve sekundi) smo dobili za vsak

posnetek tri segmente. Iz vsakega segmenta smo izračunali 1941 zvočnih značilk, med njimi statistične značilke (npr. varianco, koeficient simetrije, sploščenost), energijske značilke (npr. energije v intervalu med 250 Hz in 1 kHz), frekvenčne značilke (spektralna odstopna mesta pri 25 %, 50 %, 75 % in 90 %) in glasovne značilke (npr. tresenje, barva..., razmerje signal/šum). Celoten seznam značilk je opisan v članku Schuller in ostali [8]. Na Sliki 2 so značilke označene kot: $Z_{11}, Z_{21}, \dots, Z_{N1}$, kjer prvi indeks označuje značilko (N značilk), drugi pa segment, na katerem je bila specifična značilka izračunana. V nadaljevanju se izračuna povprečne vrednosti značilk čez vse segmente (na Sliki 2 imamo povprečje treh segmentov). Tako dobimo za vsak posnetek vektor značilk: $\text{Avg}(Z_1), \text{Avg}(Z_2), \dots, \text{Avg}(Z_N)$.

Za izbiro značilk, ki so najbolj primerne za ločevanje med obema fazama SP, smo uporabili Wilcoxonov statistični test [9], ki spada med neparametrične teste in testira, kdaj dva povezana vzorca izvirata iz iste distribucije. S tem testom ugotovimo, katere značilke imajo statistično signifikantno razliko v distribuciji glede na obe fazi.



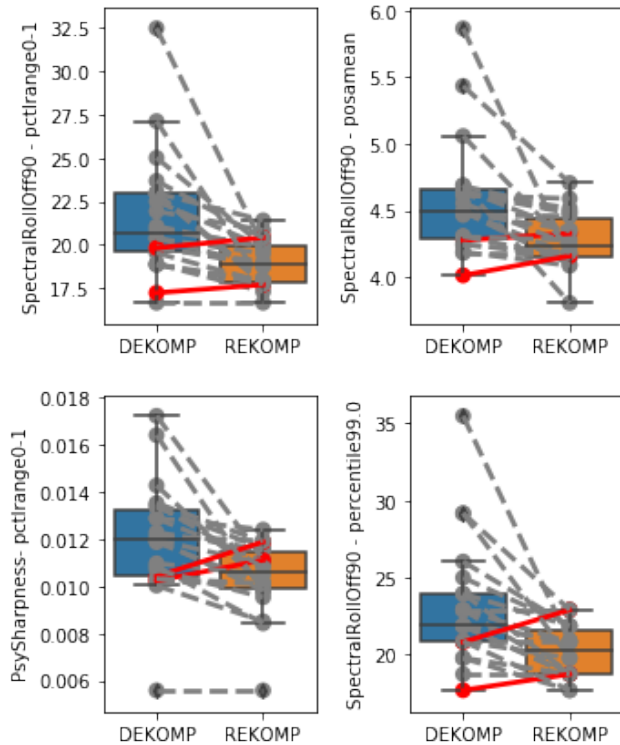
Slika 2. Shema metode

Statistične teste smo izvedli za okna velikosti: 10, 8, 6, 4, 2 in 1 sekunde s 50 % prekrivanjem. V Tabeli 1 so predstavljeni indeksi značilk, za katere je bila razlika statistično signifikantna ($p < 0.001$). Iz rezultatov lahko razberemo, da imamo za vsako velikost okna najmanj sedem informativnih značilk. Poleg tega vidimo, da število informativnih značilk narašča ob krajšanju dolžine okna. Poskusi kažejo, da je optimalna velikost okna za našo podatkovno zbirko dve sekundi s 50 % prekrivanjem.

Velikost okna (v sekundah)	10	8	6	4	2	1
# informativne značilke	7	11	14	14	16	14

Tabela 1. Število informativnih značilk, ki ločijo med rekompensirano in dekompenzirano fazo, v odvisnosti od dolžine okna.

Slika 3 predstavlja škatlaste diagrame za štiri naključno izbrane značilke izmed 16, ki smo jih dobili z uporabo okna dolžine 2 s.



Slika 3. Škatlasti diagrami za štiri značilke, ki jih statistični testi izberejo kot najbolj pomembne. Modra škatla predstavlja dekompenzirano fazo, oranžna rekompensirano. Sive črte povezujejo vrednosti značilk za posamezne paciente, kjer je vrednost v dekompenzirani fazi večja od tiste v rekompensirani, rdeče pa tiste, pri katerih velja obratno.

Štiri značilke, ki jih prikažemo v tem primeru, so sledeče:

- `fftMag_spectralRollOff90.0_sma_de_percentile1.0`
- `fftMag_spectralRollOff90.0_sma_de_percentile99.0`
- `fftMag_spectralRollOff90.0_sma_de_posamean`
- `fftMag_psySharpness_sma_de_pctrange0-1`

Vse značilke so povezane s spektralnimi lastnostmi segmenta. Prve tri značilke izračunajo Fourierovo transformacijo na kratkih oknih, izvedejo glajenje z drsečim oknom, izračunajo prvi odvod signala po času in določijo statistične označevalce, *posamean* pa predstavlja pozitivno aritmetično sredino. Značilka *psySharpness* se nanaša na spektralno strukturo signala, pri signalih, ki vsebujejo visoke frekvence, je vrednost te značilke višja. Podrobneje so te značilke opisane v [8].

Modra škatla na Sliki 3 prikazuje vrednosti značilk izračunanih iz dekompenziranih posnetkov, oranžna pa vrednosti značilk iz rekompensiranih. Siva črta ponazarja mejo, kjer je vrednost rekompensirane značilke manjša kot pripadajoča vrednost dekompenzirane značilke, rdeča črta pa ravno obratno. Iz slike je razvidno, da imamo za posamezno značilko zgolj dve rdeči črti. Če torej izvedemo binarno klasifikacijo na podlagi samo ene poljubne značilke izmed teh štirih, bomo napačno razvrstili dva izmed 22 parov posnetkov. To ustreza 91 % natančnosti metode.

Prostora za izboljšave je sicer še veliko. Načeloma bi za klasifikacijo lahko kombinirali več značilk, vendar se izkaže, da gre pri rdečih črtah večinoma za iste posameznike – to je posledica dejstva, da so značilke korelirane med sabo. Poleg tega delamo z relativno majhnim vzorcem – na 22 posameznikih, kar bi lahko pripeljalo do tega, da bi zgradili model, ki bi bil preveč prilagojen danim podatkom.

3. ZAKLJUČEK

V prispevku prikažemo prve korake h gradnji personaliziranega modela, ki bo bolnikom omogočal spremljanje poslabšanja srčnega popuščanja. Identificirali smo vrsto značilk, pri katerih obstaja statistično signifikantna razlika med vrednostmi za dekompenzirano in rekompensirano fazo. Na podlagi posamezne značilke lahko obe fazi ločimo z 91 % natančnostjo, se pa izkaže, da zaradi koreliranosti v tem trenutku ni smiselno graditi modela z uporabo več značilk.

V prihodnje nameravamo izvesti analizo na posnetkih več posameznikov, poleg tega pa bomo vključili tudi posnetke v različnih fazah, ne le v skrajnih primerih, kot v predstavljeni študiji. Če spremljamo posameznika, lahko uporabimo personaliziran model – če v določeni točki posamezne značilke, ki jih izračunamo iz zvočnih posnetkov, začnejo odstopati od pričakovanih vrednosti za rekompensirano fazo, lahko posumimo na možnost poslabšanja.

4. VIRI

- [1] Shannon M. Duplay and Véronique L. Roger. 2014. Understanding the epidemic of heart failure: past, present, and future. *Current heart failure reports* 11, 4 (2014), 404-415.
- [2] Paul A. Heidenreich, Nancy M. Albert, Larry A. Allen, David A. Bluemke, Javed Butler, Gregg C. Fonarow, John S. Ikonomidis, Olga Khavjou, Marvin A. Konstam, Thomas M. Madox, et al. 2013. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circulation: Heart Failure* 6, 3 (2013), 606-619.
- [3] William T. Abraham, Lynne W. Stevenson, Robert C. Bourge, Jo Ann Lindenfeld, Jordan G. Bauman, and Philip B. Adamson. 2016. Sustained efficacy of pulmonary artery pressure to guide adjustment of chronic heart failure therapy: complete follow-up results from the CHAMPION randomised trial. *The Lancet* 387, 10017 (2016), 453-461.
- [4] Martin Gjoreski, Monika Simjanoska, Anton Gradišek, Ana Peterlin, Matjaž Gams, and Gregor Poglajen. 2017. Chronic heart failure detection from heart sounds using a stack of machine-learning classifiers. In *2017 International Conference of Intelligent Environments (IE)*. IEEE, 14-19.
- [5] Martin Gjoreski, Anton Gradišek, Borut Budna, Matjaž Gams, and Gregor Poglajen. 2019. Toward early detection and monitoring of chronic heart failure using heart sounds.

Ambient Intelligence and Smart Environments, 26, 336-343.
DOI:10.3233/AISE190061

- [6] Martin Gjoreski, Borut Budna, Anton Gradišek, and Matjaž Gams. 2017. JSI Sound – a machine-learning tool in Orange for simple biosound classification. In *Proceedings of the Workshop on Advances in Bioinformatics and Artificial Intelligence: Bridging the Gap, co-located with 26th International Joint Conference on Artificial Intelligence (IJCAI 2017)*. Melbourne, Australia, 26-29.
- [7] Florian Eyben, Felix Weninger, Florian Gross, and Björn Schuller. 2013. Recent developments in OpenSMILE, the munich open-source multimedia feature extractor. In *Proceedings of the 21st ACM international conference on Multimedia*. ACM, 835-838.
- [8] Björn Schuller, Michel Valstar, Florian Eyben, Gary McKenown, Roddy Cowie, and Maja Pantic. 2011. Avec 2011-the first international audio/visual emotion challenge. In *International Conference of Affective Computing and Intelligent Interaction*. Springer, 415-424.
- [9] Frank Wilcoxon. 1992. Individual comparison by ranking methods. In *Breakthroughs in statistics*. Springer, 196-202.