



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA STROJNÍHO INŽENÝRSTVÍ**

FACULTY OF MECHANICAL ENGINEERING

**ÚSTAV AUTOMATIZACE A INFORMATIKY**

INSTITUTE OF AUTOMATION AND COMPUTER SCIENCE

**KOLABORACE ČLOVĚK STROJ – VYUŽITÍ  
ZPRACOVÁNÍ ŘEČI**

HUMAN-MACHINE COLLABORATION – USING SPEECH PROCESSING

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

**Štěpán Kisler**

**VEDOUCÍ PRÁCE**

SUPERVISOR

**Ing. Martin Juříček**

**BRNO 2024**



# Zadání bakalářské práce

Ústav: Ústav automatizace a informatiky  
Student: **Štěpán Kisler**  
Studijní program: Strojírenství  
Studijní obor: Aplikovaná informatika a řízení  
Vedoucí práce: **Ing. Martin Juříček**  
Akademický rok: 2023/24

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma bakalářské práce:

## Kolaborace člověk–stroj – využití zpracování řeči

### Stručná charakteristika problematiky úkolu:

Teoretická část práce bude zahrnovat rešerši o kolaborativních robotech ze stáje Universal Robots a analýzu metod pro řešení zpracování řeči. Hlavní náplní práce bude využití metod pro zpracování řeči a využití pro řízení kolaborativního robota UR3. Práce taktéž bude zahrnovat ověření řešení jak v simulaci, tak v laboratoři. Práce předpokládá aktivní přístup. Využít lze programovací jazyk Python nebo Matlab.

### Cíle bakalářské práce:

- Provedte rešerši kolaborativních robotů společnosti Universal Robots a blíže popište zvoleného kolaborativního robota UR3
- Provedte rešerši o metodách zpracování řeči
- Zvolte a implementujte metodu pro zpracování řeči
- Ověřte funkčnost vytvořeného řešení pomocí simulace a v laboratoři

### Seznam doporučené literatury:

SICILIANO, Bruno a KHATIB, Oussama, ed. Springer handbook of robotics. 2nd edition. Berlin: Springer, [2016]. ISBN 978-3-319-32550-7.

KOLÍBAL, Zdeněk. Roboty a robotizované výrobní technologie. Brno: Vysoké učení technické v Brně - nakladatelství VUTIUM, 2016. ISBN 978-80-214-4828-5

Universal Robots: Collaborative Robotics. <https://www.universal-robots.com/>

GOLD, Ben, Nelson MORGAN a Dan ELLIS. Speech and Audio Signal Processing: Processing and Perception of Speech and Music. 2nd Edition. Wiley, 2011. ISBN 978-0-470-19536-9.

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2023/24

V Brně, dne

L. S.

---

ředitel ústavu

---

doc. Ing. Jiří Hlinka, Ph.D.  
děkan fakulty

## **ABSTRAKT**

Tato bakalářská práce se zabývá návrhem a implementací systému hlasového ovládání pro kolaborativního robota UR3 CB série od společnosti Universal Robots, s cílem zjednodušit interakci mezi člověkem a robotem. V úvodu práce je představena kolaborativní robotika, její historie, příklady úspěšných aplikací a možnosti programování kolaborativních robotů. Dále se práce věnuje technologii rozpoznávání řeči, jejímu využití, historii a metodám. Praktická část zahrnuje porovnání existujících systémů pro rozpoznávání řeči a výběr nejvhodnějšího z nich pro hlasové ovládání robota. Popisuje také návrh programu pro hlasové ovládání v jazyce Python a testování celého systému v simulaci i v reálných podmínkách v robotické laboratoři.

## **ABSTRACT**

This bachelor's thesis focuses on the design and implementation of a voice control system for the UR3 CB series collaborative robot from Universal Robots, aiming to simplify human-robot interaction. The introduction provides an overview of collaborative robotics, including its history, successful applications, and the possibilities of programming collaborative robots. Additionally, it explores speech recognition technology, covering its applications, history, and methods. The practical section compares existing speech recognition systems and selects the most suitable one for robot voice control. It also details the development of a voice control program in Python and the testing of the entire system, both in simulation and real-world conditions in a robotics laboratory.

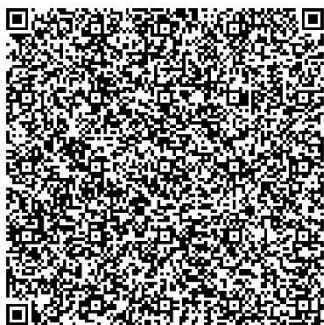
## **KLÍČOVÁ SLOVA**

Kolaborativní robotika, rozpoznávání řeči, hlasové ovládání, Universal Robots, hluboké neuronové sítě, Whisper, IBM Watson Speech To Text, Google Speech Recognition, UR3

## **KEYWORDS**

Collaborative robotics, speech recognition, voice control, Universal Robots, deep neural networks, Whisper, IBM Watson Speech to Text, Google Speech Recognition, UR3





2024

## **BIBLIOGRAFICKÁ CITACE**

KISLER, Štěpán. Kolaborace člověk-stroj – využití zpracování řeči. Brno, 2024. Dostupné také z: <https://www.vut.cz/studenti/zav-prace/detail/157234>. Bakalářská práce. Vysoké učení technické v Brně, Fakulta strojního inženýrství, Ústav automatizace a informatiky. Vedoucí práce Martin Juříček.





## **PODĚKOVÁNÍ**

Tímto bych chtěl poděkovat svému vedoucímu práce Ing. Martinovi Juříčkovi a konzultantovi Ing. Adamovi Vargovi za jejich cenné rady, trpělivost a podporu, kterou mi poskytovali během zpracování mé práce. Zároveň děkuji své rodině za neustálou podporu během mého studia a v neposlední řadě své přítelkyni za podporu a pomoc s jazykovou korekturou.



## ČESTNÉ PROHLÁŠENÍ

Prohlašuji, že tato práce je mým původním dílem, vypracoval jsem ji samostatně pod vedením vedoucího práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury.

Jako autor uvedené práce dále prohlašuji, že v souvislosti s vytvořením této práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následku porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestně právních důsledků.

V Brně dne 20. 5. 2024

.....

Štěpán Kisler



# OBSAH

<b>1</b>	<b>ÚVOD.....</b>	<b>15</b>
<b>2</b>	<b>KOLABORATIVNÍ ROBOTIKA.....</b>	<b>17</b>
2.1	Definice kolaborativních robotů a jejich vlastnosti .....	17
2.2	Historie a vývoj kolaborativní robotiky.....	18
2.3	Aplikace kolaborativních robotů .....	20
2.4	Programování a ovládání kolaborativních robotů .....	22
2.5	Universal Robots .....	24
2.5.1	Kolaborativní robot UR3 CB série .....	26
<b>3</b>	<b>ROZPOZNÁVÁNÍ ŘEČI.....</b>	<b>27</b>
3.1	Definice a současné využití rozpoznávání řeči.....	27
3.2	Historie a vývoj rozpoznávání řeči .....	28
3.3	Teoretický základ: HMM a DNN.....	29
3.3.1	Skryté Markovovy modely (HMM) .....	29
3.3.2	Hluboké neuronové sítě (DNN).....	30
3.4	Metody rozpoznávání řeči .....	32
3.4.1	Tradiční systémy.....	32
3.4.2	End-to-end systémy .....	32
3.4.3	Porovnání tradičních a end-to-end systémů.....	33
<b>4</b>	<b>PRAKTICKÁ ČÁST .....</b>	<b>35</b>
4.1	Výběr vhodného nástroje pro rozpoznávání řeči .....	35
4.1.1	Whisper.....	35
4.1.2	IBM Watson Speech to Text .....	36
4.1.3	Google Speech Recognition prostřednictvím knihovny SpeechRecognition... ..	38
4.1.4	Porovnání jednotlivých systémů.....	39
4.2	Návrh programu pro hlasové ovládání kolaborativního robota.....	43
4.2.1	Požadavky na program .....	43
4.2.2	Volba nástrojů.....	44
4.2.3	Funkce a architektura navrženého programu.....	44
4.3	Testování funkčnosti řešení pomocí simulace a v laboratoři .....	47
<b>5</b>	<b>DISKUZE .....</b>	<b>51</b>
<b>6</b>	<b>ZÁVĚR .....</b>	<b>53</b>
	<b>SEZNAM POUŽITÉ LITERATURY.....</b>	<b>55</b>
	<b>SEZNAM ZKRATEK A SYMBOLŮ .....</b>	<b>59</b>
	<b>SEZNAM OBRÁZKŮ .....</b>	<b>61</b>
	<b>SEZNAM TABULEK.....</b>	<b>63</b>
	<b>SEZNAM PŘÍLOH.....</b>	<b>65</b>



# 1 ÚVOD

Budoucnost nespočívá v nahrazování lidí roboty, nýbrž v rozvíjení nových způsobů vzájemné spolupráce. Tím by tato technologie osvobodila lidi od rutinních úkolů a umožnila jim soustředit se čistě na lidské dovednosti, jako je kreativní myšlení, rozhodování a sociální interakce. Za tímto účelem byli vyvinuti kolaborativní roboti, kteří jsou schopni bezpečně pracovat po boku lidí v široké škále aplikací, od výroby a montáže až po zdravotnictví a osobní asistenci.

Jedním z klíčových aspektů této spolupráce je efektivní komunikace, která by měla být co nejpřirozenější, a tou je pro lidi nepochybně mluvená řeč. To vede k myšlence hlasového ovládání kolaborativních robotů, kterou se zabývá tato bakalářská práce. Kromě zvýšení komfortu uživatelů tato technologie nabízí značný potenciál pro zefektivnění pracovních procesů. Příkladem může být využití hlasového ovládání na výrobních linkách, jenž by operátorům umožnilo udržovat plynulý pohyb a soustředit se současně na více úkolů bez potřeby fyzické interakce s ovládacím panelem. Další možností neprůmyslového využití by mohla být asistence osobám s postižením pohybového aparátu, která by jim poskytla větší míru samostatnosti a snížila jejich závislost na pomoci druhých.

Hlavním cílem této bakalářské práce je návrh systému hlasového ovládání pro kolaborativního robota UR3 od společnosti Universal Robots. Úvodní část práce představuje kolaborativní robotiku, její historii, příklady aplikací kolaborativních robotů v různých odvětvích a možnosti jejich programování. Dále je popsána společnost Universal Robots a jejich model UR3 CB série. Následující část se věnuje technologii rozpoznávání řeči, jejímu využití, historii a metodám. Praktická část obsahuje porovnání existujících systémů pro rozpoznávání řeči a výběr nejvhodnějšího z nich. Následně je navržen program pro hlasové ovládání v programovacím jazyce Python. Na závěr je provedeno testování navrženého programu nejprve v simulaci a následně na reálném robotovi v robotické laboratoři na Ústavu automatizace a informatiky.





## 2 KOLABORATIVNÍ ROBOTIKA

### 2.1 Definice kolaborativních robotů a jejich vlastnosti

Kolaborativní robot, často označován jako kobot, je typ průmyslového robota navržený tak, aby mohl bezpečně a efektivně spolupracovat s lidmi při různých činnostech. Od běžných robotů se liší především tím, že nemusí být izolován od lidí ochranným oplocením nebo bariérami, ale může s nimi pracovat v těsné blízkosti či dokonce v přímém kontaktu. Mezi hlavní vlastnosti kolaborativních robotů patří především:

- **Bezpečnost**, která představuje nezpochybnitelně jejich nejdůležitější vlastnost. Koboti jsou vybaveni širokým spektrem bezpečnostních senzorů a jejich rychlosti a síly jsou omezeny tak, aby při kontaktu s člověkem nemohli způsobit žádná zranění. Důležité je však zdůraznit, že ačkoliv jsou určeni k práci v blízkosti lidí, nemohou k tomu být využiti ve všech situacích. Pokud kobot manipuluje s nebezpečnými nástroji nebo předměty, nelze mu umožnit přímý kontakt s obsluhou. Z tohoto důvodu se musí z pohledu bezpečnosti posuzovat celý systém či aplikace. Spolupráce mezi člověkem a kobotem je možná díky normám ISO 10218-1/-2 a doplňkové technické specifikaci ISO/TS 15066, přičemž společně se zabývají problémy bezpečnosti kolaborativních robotů. Technická specifikace říká, že při potenciálním kontaktu člověka a kobota nesmí dojít k jakémukoliv zranění či způsobení bolesti, a současně definuje možnosti a bezpečnostní limity takové spolupráce [1].
- **Flexibilita**, díky ní je lze snadno přesouvat a přizpůsobovat různým pracovním úkolům a prostředím. Koboti jsou schopni vykonávat určitou činnost během jedné pracovní směny, po které mohou být přemístěni na jiné pracoviště, kde se budou věnovat odlišným činnostem v následující směně.
- **Snadná programovatelnost**, ta umožňuje, že i lidé bez pokročilejších znalostí programování mohou koboty ovládat pomocí intuitivních rozhraní.

## 2.2 Historie a vývoj kolaborativní robotiky

Před 103 lety proběhlo v českém Národním divadle první představení hry Karla Čapka s názvem R.U.R (Rossumovi univerzální roboti). V této hře poprvé zaznělo světoznámé slovo *robot*, které vymyslel autorův bratr Josef Čapek. Tímto termínem v divadelní hře nazval ony umělé dělníky, kteří měli sloužit lidem. Od té doby uběhlo více než 30 let do vynálezu prvního opravdového robota s označením Programmed Article Transfer, za kterým stál americký vynálezce George Devol. Ten se později setkal s podnikatelem Josephem Engelbergem a společně vytvořili prototyp robota se jménem Unimate #001. Vážil přibližně 1400 kg a svým vzhledem více připomínal paži než lidskou postavu (viz obr. 1). V roce 1959 byl poprvé využit v továrně General Motors na lince pro tlakové lití a v roce 1961 se stal prvním sériově vyráběným průmyslovým robotem. Později se společnost General Motors s roboty Unimate stala nejautomatizovanější automobilkou na světě a porážela konkurenci. Ostatní automobilky se jí v tomto směru snažily následovat, a to přispělo vzniku mnoha dalších společností zajímajících se o robotiku [2; 3].



Obr. 1: Robot Unimate [42]

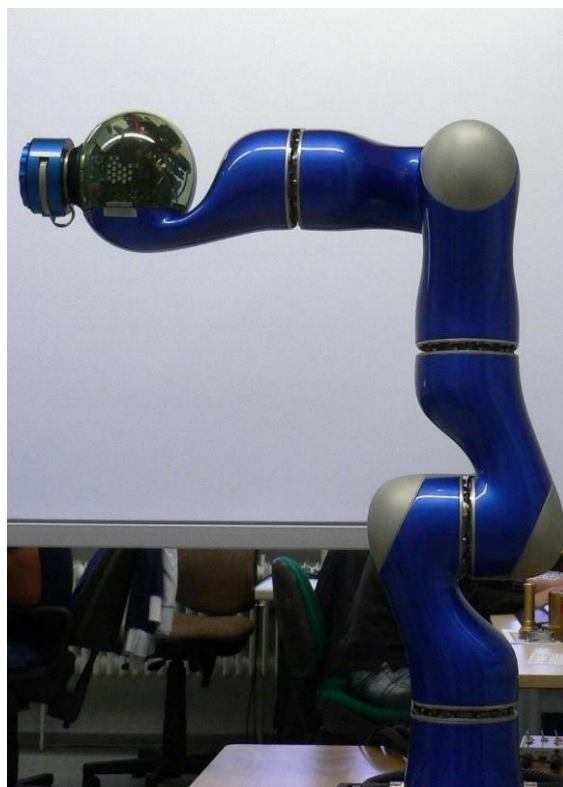
V devadesátých letech 20. století se Úřad pro bezpečnost a ochranu zdraví při práci (OSHA) v USA začal věnovat ergonomii v automobilových závodech s obavami o zdraví zaměstnanců. Monotónní a silově náročná práce výrazně zatěžovala jejich těla, což v průběhu času vedlo ke zdravotním problémům. Zavedením mechanických a automatizovaných řešení mohl být tento problém eliminován. V reakci na to se společnost General Motors obrátila na Northwestern University a tato spolupráce dala vzniknout prvním formám kolaborativních robotů, které v roce 1996 vynalezli profesori J. Edward Colgate a Michael Peshkin. Jejich patent, pod názvem *Cobots*, popisuje „aparát a metodu pro přímou fyzickou interakci mezi člověkem a všeobecným manipulátorem řízeným počítačem“ [4]. Společně založili v roce 1997 společnost Cobotics, jež vyráběla

koboty určené především k manipulaci s díly při montáži automobilů (obr. 2). Tito koboti neměli z bezpečnostních důvodů žádnou vnitřní hnací sílu, která by jim umožňovala pohyb nezávislý na externí síle, kterou vyvíjel samotný člověk. Podle síly a tlaku vyvíjeného pracovníkem se dokázali přizpůsobovat jeho pohybům [5].



Obr. 2: Kobot od společnosti Cobotics [43]

Za prvního zástupce blížícím se moderním kolaborativním robotům by se dal považovat lehký robot LWR III vydaný společností KUKA v roce 2004 (obr. 3). Tento robot vznikl po dlouholeté spolupráci společnosti KUKA s Německým centrem pro letectví a kosmonautiku [6]. Měl 7 stupňů volnosti a jeho poměr mezi vlastní vahou a maximální nosností byl 1:1 s hodnotou 14 kg. S takto nízkou hmotností a momentovými senzory na každém kloubu byl schopný interakce s člověkem [7]. V roce 2008 vydala dánská společnost Universal Robots prvního komerčního kolaborativního robota s označením UR5, který byl nainstalován téhož roku v dánské společnosti Linatex, kde pracoval bezpečně po boku lidí. Zaměstnanci si ho dokázali sami naprogramovat i přesto, že sami neměli žádné hlubší znalosti programování. Tímto společnost Universal Robots udala směr dalšímu vývoji flexibilních, snadno programovatelných kobotů, u kterých není nutností klecové ohrazení a dodnes je lídrem na trhu [8].



Obr. 3: Robot KUKA LWR III [44]

V současnosti je kolaborativní robotika nejrychleji rostoucím odvětvím v oblasti robotického průmyslu. K roku 2023 měl trh s kolaborativními roboty hodnotu 1,2 miliardy dolarů a předpokládá se, že do roku 2029 dosáhne zhruba 6,8 miliardy dolarů [9]. Mezi největší výrobce na trhu s kolaborativními roboty patří kromě Universal Robots jmenovitě také FANUC, ABB nebo KUKA.

### 2.3 Aplikace kolaborativních robotů

V současnosti se koboti, díky svým vlastnostem, stávají klíčovým nástrojem pro automatizaci v mnoha odvětvích. Největší využití však mají stále ve výrobním průmyslu, kde zastávají především monotónní práce jako například přemísťování předmětů z bodu A do bodu B (pick and place), balení a skládání na palety, svařování, kontrolu kvality výrobků pomocí strojového vidění nebo různé montážní operace [10].

Na svém inovativním montážním pracovišti využívá koboty i Fraunhofer IOSB ve spolupráci s divizí PwC Strategy& (viz obr. 4). Hlavním cílem tohoto pracoviště je umožnit efektivní a bezpečnou spolupráci mezi pracovníkem a kobotem při ruční montáži. Tento systém pracovníkovi promítá instrukce na pracovní plochu, a díky kombinaci metod strojového učení se dokáže přizpůsobit jeho pohybům. Umožňuje pracovníkovi komunikovat prostřednictvím gest či hlasových příkazů, které jsou velmi užitečné v situacích, kdy pracovník nemá volné ruce [11].



Obr. 4: Chytré montážní pracoviště CoBot [11]

Své místo zaujímají kolaborativní roboti i ve zdravotnictví, kde přispívají ke zvýšení efektivity a umožňují kvalifikovaným pracovníkům se zaměřit na důležitější úkoly. Příkladem může být aplikace dvojramenného kobota YuMi od společnosti ABB v laboratoři univerzitní nemocnice Karolinska ve Švédsku (obr. 5). Kobota zde využívají ke skenování zkumavek a otevírání transportních obalů, což do té doby výrazně zatěžovalo zápěstí zaměstnanců i jejich čas [12]. Tento kobot se nachází také v robotické laboratoři na ÚAI.



Obr. 5: Kolaborativní robot YuMi od společnosti ABB manipulující se zkumavkami v univerzitní nemocnici Karolinska [12]

Obrovské uplatnění mají koboti i v potravinářském průmyslu a zahradnictví, kde se využívají například k balení potravin, zdobení dortů nebo v nových robotických systémech pro sběr ovoce a zeleniny. Příkladem takového systému je brněnský start-up FRAVEBOT (FRuit And VEgetable roBOT). Tento robot složený z jednoho nebo dvou kobotů od společnosti Universal Robots na mobilní platformě, je využíván ke sběru jahod a k monitorování zdravotního stavu rostlin a plodů (obr. 6) [13].



Obr. 6: FRAVEBOT při sběru jahod [13]

## 2.4 Programování a ovládání kolaborativních robotů

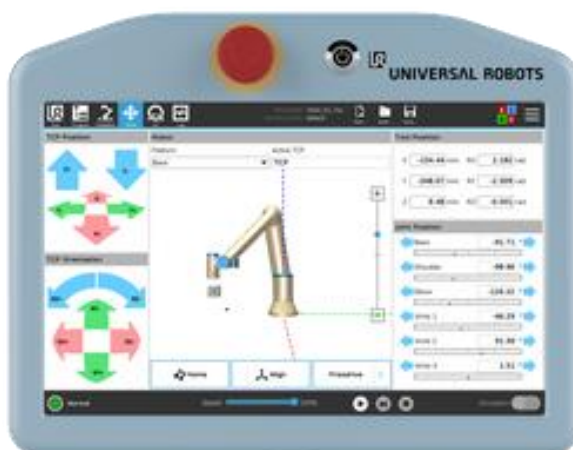
V odvětví kolaborativní robotiky je kladen důraz na to, aby bylo možné koboty ovládat a programovat co nejjednodušeji a nejintuitivněji, což, jak již bylo zmíněno, představuje jednu z jejich největších předností. Díky tomu je mohou využívat ve svých provozech i lidé, kteří nejsou příliš zdatní v programování bez nutnosti najímání profesionálů či investic do nákladných školení. Tento přístup výrazně ulehčuje malým nebo středním podnikům vstup do světa automatizace. V této kapitole bude představeno stručné shrnutí možností programování kobotů od té nižší úrovně s využitím programovacích jazyků od výrobců až po ovládání skrze interakci založené na umělé inteligenci.

Za programování kobotů na základní úrovni může být považováno použití speciálních programovacích jazyků vyvinutých přímo výrobcem těchto zařízení. Tyto jazyky jsou často navrženy tak, aby co nejlépe vyhovovaly hardwarovým a softwarovým specifikacím daného robota a umožňovaly uživateli přesně definovat jeho pohyby a operace. Mezi tyto programovací jazyky patří například KRL (KUKA Robot Language) od KUKA, RAPID od ABB, URScript od Universal Robots nebo INFORM od Yaskawa. Kromě toho někteří výrobci také nabízejí své vlastní vývojové prostředí, jako například ABB RobotStudio nebo FANUC RoboGuide, které poskytují uživatelům komplexní nástroje pro programování, simulaci a vizualizaci robotických operací. Tato prostředí

výrazně zjednodušují testování a optimalizaci aplikací před jejich nasazením v reálném provozu.

Přecházení mezi jazyky výrobců však nebylo příliš komfortní, a proto přinesla integrace obecných programovacích jazyků jako je C++ nebo Python, prostřednictvím rozhraní API<sup>1</sup> nebo díky rozvoji platforem jako ROS<sup>2</sup> příjemnou změnu. Programování robotů je s těmito jazyky mnohem flexibilnější, zároveň zjednodušuje integraci robotů s jinými technologiemi, jako jsou různé senzory nebo IoT<sup>3</sup>.

Dalším nástrojem pro programování kobotů je kompaktní ovládací panel, známý jako Teach Pendant, který je standardní součástí vybavení většiny kobotů (obr. 7). Díky svému intuitivnímu uživatelskému rozhraní umožňuje rychlé a snadné programování. S jeho pomocí je možné kobotu „učit“ prostřednictvím přímé demonstrace, kdy ho člověk vede ručně nebo pomocí ovládacího panelu po určité trajektorii a v průběhu mu přiřazuje činnosti, které má v danou chvíli vykonat. Kobot si tuto sekvenci zapamatuje a poté ji dokáže kdykoliv zopakovat. Někteří výrobci navíc na ovládacím panelu umožňují využití vizuálních programovacích jazyků, pomocí kterých lze robota programovat jednoduše přetahováním bloků či ikon, reprezentujících různé funkce.



Obr. 7: Teach Pendant s uživatelským rozhraním Polyscope 5 od společnosti Universal Robots [45]

Koboty lze ovládat také pro člověka přirozenějším způsobem, a to prostřednictvím verbální (hlasové pokyny) a neverbální (gesta, pózy, haptika) komunikace. V první řadě jsou naprogramovány různé pohyby a činnosti, které se následně, verbálním nebo neverbálním příkazem spouští [14]. Klíčovým prvkem, jenž tuto interakci umožňuje a zefektivňuje, je využití umělé inteligence, která stojí za přesným rozpoznáním

<sup>1</sup> API (Application Programming Interface) je soubor pravidel, který umožňuje, aby různé programy spolupracovaly a vzájemně si vyměňovaly informace.

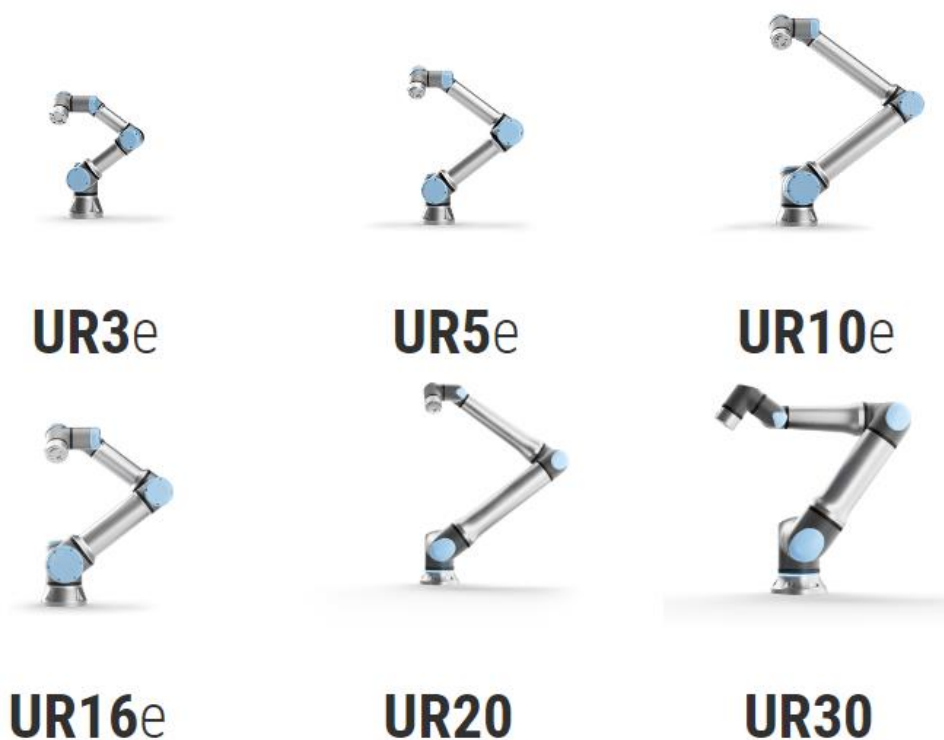
<sup>2</sup> ROS (Robot Operating System) je komplexní open-source framework pro tvorbu softwaru pro roboty.

<sup>3</sup> IoT (Internet of Things) je síť vzájemně propojených fyzických zařízení, které spolu komunikují a vyměňují data skrze internet, díky čemuž je možné je na dálku ovládat a monitorovat.

a interpretací uživatelských pokynů. Tato metoda bude podrobněji popsána v dalších kapitolách.

## 2.5 Universal Robots

Dánskou společností Universal Robots, specializující se na výrobu kolaborativních robotů, založili v roce 2005 Esben Østergaard, Kristian Kassow a Kasper Støy, tehdejší studenti University of Southern Denmark sídlící ve městě Odense. Myšlenka jejich projektu spočívala ve zpřístupnění robotizace menším společnostem, což se díky investici společnosti Syddansk Innovation povedlo. Jak již bylo zmíněno v kapitole 2.2, rok 2008 představoval obrovský milník nejen pro společnost Universal Robots, ale také pro celou robotickou komunitu. V tomto roce se jim podařilo prodat prvního komerčního kobotu na světě, který nesl označení UR5. V roce 2012 byl představen další model UR10 s větší nosností a dosahem. O tři roky později, v roce 2015, představili i menší stolní verzi UR3, čímž byla kompletní první řada s označením CB. Téhož roku došlo k odkoupení Universal Robots americkou společností Teradyne za 285 milionů dolarů. V roce 2018 byla představena nová řada s názvem „e-Series“ s modely UR3e, UR5e a UR10e. Nabízela mnohá vylepšení, jako například nové bezpečnostní funkce nebo větší přesnost a citlivost, což jim otevřelo dveře k novým aplikacím. Tuto řadu o rok později rozšířili o model UR16e s větší nosností. Nejnovějšími modely jsou aktuálně UR20 a UR30, které opět posunuly hranice kobotů od společnosti Universal Robots. Všechny aktuální modely jsou na (obr. 8) a jejich parametry uvedeny v tab. 1 [15; 16].



Obr. 8: Přehled všech nabízených kolaborativních robotů od společnosti Universal Robots [46]



Tab. 1: Parametry všech nabízených kolaborativních robotů od společnosti Universal Robots [17]

Parametry	Model					
	UR3e	UR5e	UR10e	UR16e	UR20	UR30
Nosnost [kg]	3	5	12,5	16	20	30
Dosah [mm]	500	850	1300	900	1750	1300
Opakovatelnost [mm]	±0,03	±0,03	±0,05	±0,05	±0,1	±0,1
Rychlost [m/s]	1	1	1	1	2	2
Hlučnost [dB]	<60	<65	<65	<65	<65	<65
Hmotnost [kg]	11,2	20,6	33,5	33,1	64	63,5

Společnost zároveň nabízí svůj ekosystém UR+. Jedná se o certifikované příslušenství a doplňky třetích stran, které rozšiřují možnosti a funkcionalitu kobotů Universal Robots a zjednodušují jejich implementaci do různých procesů. Mezi tyto produkty patří například různé typy gripperů,<sup>4</sup> senzorické systémy, systémy počítačového vidění, software, bezpečnostní vybavení a další. UR+ poskytuje partnerům vyrábějícím tyto doplňky přístup k širšímu trhu a svým zákazníkům zajišťuje záruku kompatibility [18]. Obr. 9 znázorňuje paletizační řešení řady AX od společnosti ROBOTIQ, které slouží ke skládání výrobků z linky na palety. Řada AX má integrovanou vertikální osu, pomocí níž může skládat předměty až do výšky 2,75 metru [19].



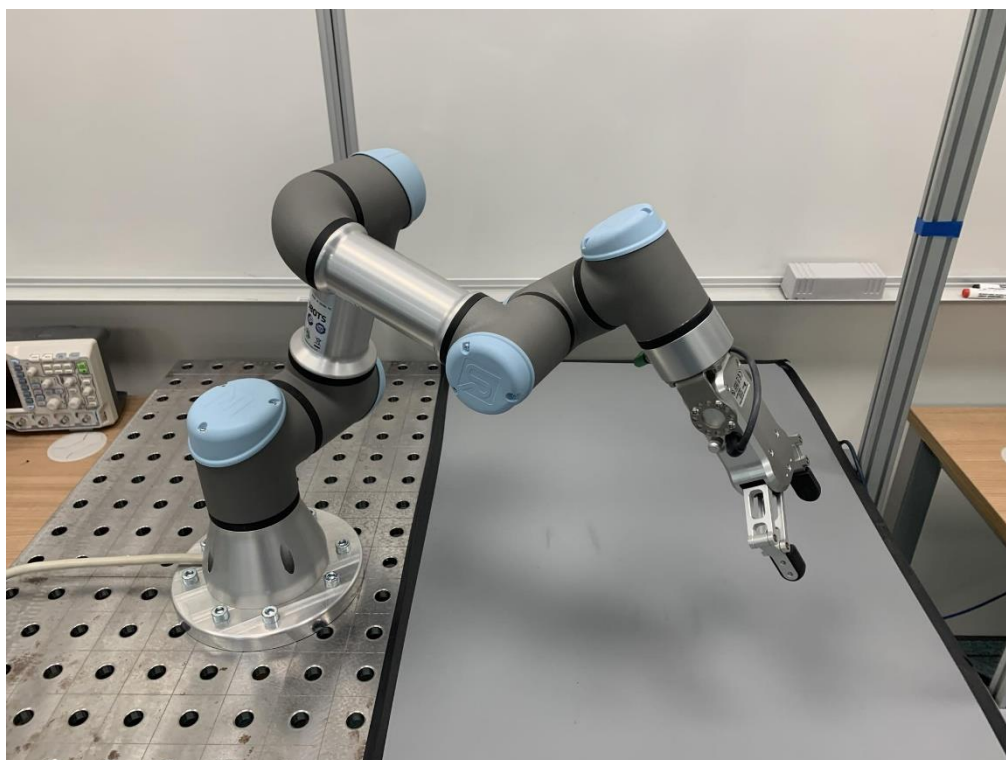
Obr. 9: Paletizační řešení od společnosti ROBOTIQ [19]

<sup>4</sup> Gripper je koncové zařízení robotického ramene, které umožňuje robotu uchopit objekty a manipulovat s nimi. Mezi typy gripperů patří mechanické kleště, sací plochy a další specializované adaptéry.

### 2.5.1 Kolaborativní robot UR3 CB série

Kolaborativní robot UR3 z řady CB je k dispozici v robotické laboratoři Ústavu automatizace a informatiky (viz obr. 10) a bude využit v rámci této bakalářské práce. Je zde detailně popsán právě z důvodu jeho významné role při řešení tohoto úkolu.

Jedná se o nejmenší model z nabídky Universal Robots, který je díky svým malým rozměrům vhodný k instalaci na pracovní stoly nebo k integraci do různých zařízení. Tento hliníkový kobot s hmotností pouhých 11 kg má nosnost 3 kg a dosah až 500 mm při opakovatelnosti s přesností  $\pm 0,1$  mm. Může se pohybovat v šesti osách, přičemž s každým kloubem může rotovat o  $360^\circ$  a s koncovým kloubem neomezeně. Pro správnou funkci musí být okolní teplota v rozmezí  $0-50^\circ\text{C}$ . Disponuje poměrně nízkou hlučností s hodnotou okolo 70 dB a je opatřen stupněm krytí IP 64, což znamená, že je chráněn proti vniknutí prachu a stříkající vody ze všech směrů. Je vybaven čtyřmi vstupy, z nichž dva jsou digitální a dva analogové, a také dvěma analogovými výstupy. Je k němu rovněž dodán dotykový ovládací panel s uživatelským rozhraním UR Polyscope a kontrolér obsahující 16 digitálních vstupů a výstupů a dva analogové vstupy a výstupy. Komunikace může probíhat prostřednictvím protokolů TCP/IP, Modbus TCP, Profinet nebo Ethernet IP. Odhadovaná životnost se při jmenovitém zatížení udává 35 000 hodin [20].



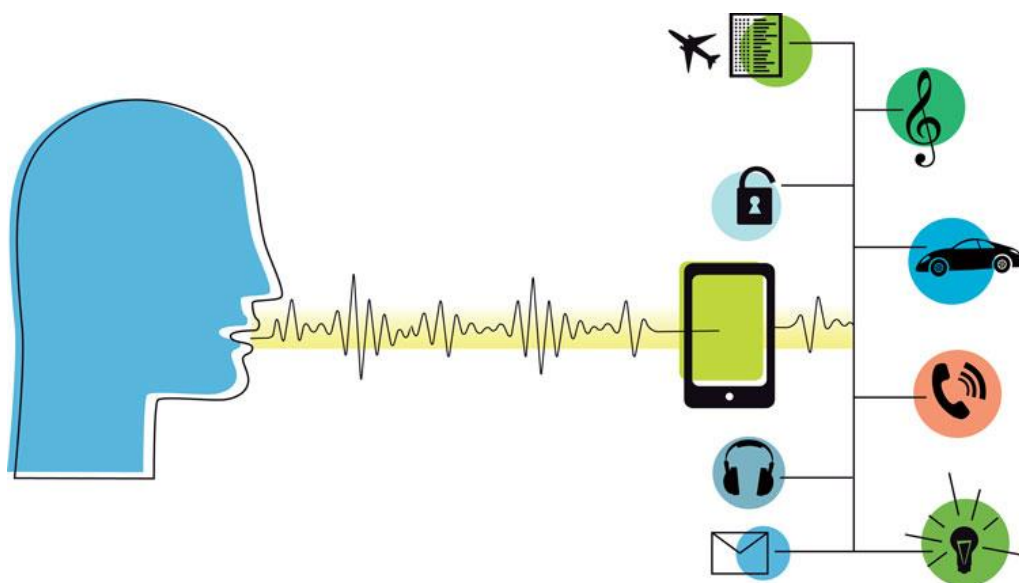
Obr. 10: Kolaborativní robot UR3 CB série v robotické laboratoři na ÚAI

## 3 ROZPOZNÁVÁNÍ ŘEČI

### 3.1 Definice a současné využití rozpoznávání řeči

Rozpoznávání řeči představuje technologii, jež umožňuje počítačům rozpoznávat a zpracovávat mluvená slova a převádět je do textové formy. Mluvená řeč je druh akustického signálu, konkrétně se jedná o jednorozměrný (1D) signál. Hlas je pro člověka jednou z nejpřirozenějších forem komunikace, z toho důvodu hraje rozpoznávání řeči významnou roli v rámci interakce mezi člověkem a strojem. Pomocí této technologie můžeme hlasovými příkazy ovládat různá zařízení, aplikace nebo služby bez nutnosti fyzického kontaktu [21].

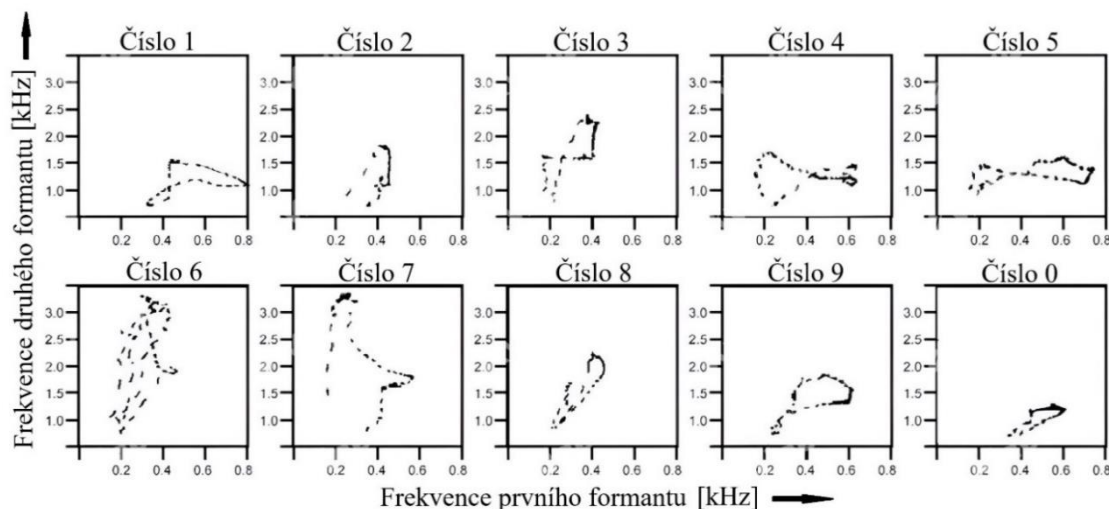
V současnosti se tato technologie nejvíce využívá u virtuálních hlasových asistentů, jako je Siri od společnosti Apple, Alexa od Amazonu či Google Assistant a další. S jejich pomocí je možné ovládat prostřednictvím hlasových příkazů různé funkce na mobilním telefonu, vyhledávat informace na internetu nebo ovládat chytrou domácnost. Hlasové ovládání se významně prosazuje i v automobilovém průmyslu, kde nabízí řidičům možnost v průběhu jízdy ovládat navigační systémy, vyřizovat telefonní hovory či ovládat hudbu, aniž by museli odvracet zrak od silnice před sebou. Tato technologie tak přispívá ke zvýšení bezpečnosti a komfortu během řízení. V oblasti logistiky a skladování zjednodušuje komunikaci s řídicími systémy, umožňuje efektivnější řízení zásob a snižuje chybovost při vychystávání objednávek. Ve zdravotnictví mohou lékaři využít rozpoznávání řeči pro efektivní zaznamenávání poznámek o diagnóze pacientů. Tato činnost je považována za jednu z časově nejnáročnějších v tomto prostředí. Díky tomu se mohou lékaři věnovat dalším pacientům [22]. Tyto příklady ukazují jasný potenciál technologie rozpoznávání řeči a její pozitivní dopad na mnohá odvětví.



Obr. 11: Ilustrace využití rozpoznávání řeči [47]

### 3.2 Historie a vývoj rozpoznávání řeči

Vývoj rozpoznávání řeči, probíhající již více než sedm desetiletí, odhaluje pozoruhodnou cestu od prvních experimentálních systémů rozpoznávajících omezenou sadu slov, po moderní sofistikované metody založené na algoritmech hlubokého učení. Za první milník lze považovat systém *Audrey*, vyvinutý v roce 1952 v Bell Laboratories, který fungoval na principu analýzy frekvence formantů,<sup>5</sup> které byly měřeny během vyslovení samohlásek v každé číslici (viz obr. 12). Systém dokázal rozpoznávat čísla od nuly do devíti vyřčené jedním mluvčím [23].



Obr. 12: Závislost formantu 1 vůči formantu 2 reprezentující jednotlivá čísla (popisky přeloženy autorem) [23]

O 10 let později vyvinula společnost IBM systém *Shoebbox*, pojmenovaný podle jeho malé velikosti odpovídající tehdejší běžné americké krabici na obuv. Shoebbox dokázal rozpoznat 16 slov a byl schopen provádět základní matematické operace [24]. Dalším důležitým milníkem byl systém *Harpy*, který vznikl v rámci programu Speech Understanding Research (SUR), jenž byl iniciativou Agentury pro pokročilé obranné výzkumné projekty (DARPA). Tento projekt financovalo americké ministerstvo obrany mezi lety 1971 až 1976. Harpy byl schopný rozpoznat až 1011 slov, což představovalo významný pokrok.

V 80. letech 20. století došlo k výraznému zlomu, kapacita systémů pro rozpoznávání řeči se dramaticky rozšířila z několika stovek slov na několik tisíc. Tento pokrok umožnilo rozšíření použití statistických metod, zejména skrytých Markovových modelů (HMM) [25]. Významný potenciál statistického přístupu ukázal v polovině 80. let systém od společnosti IBM zvaný *Tangora*, který vyvinul Frederick Jelínek,

<sup>5</sup> Formanty jsou vrcholové frekvence ve zvukovém spektru hlasu, ovlivněné rezonancemi hlasového traktu. Jsou klíčové pro identifikaci samohlásek v řeči, jelikož první formant naznačuje, do jaké míry jsou naše ústa otevřená nebo zavřená při jejich artikulaci. S otevřenějšími ústy, jako při vyslovnosti samohlásky „a“, je formant nižší. Druhý formant pomáhá určit, kde v ústech tvoříme zvuk, jestliže blíže k přední části, jako u samohlásky „i“, nebo blíže k zadní části, jako u „u“ [41].

Čechoslovák žijící v USA, s jeho týmem. Tento systém zvládl rozpoznat kontinuální řeč se slovníkem o rozsahu až 20 000 slov [26]. V 90. letech přinesl pokrok v mikroprocesorové technologii první komerčně dostupný desktopový systém rozpoznávání řeči, *Dragon Dictate*, ten byl později vylepšen a nazván *Dragon NaturallySpeaking*, se schopností rozpoznávat až 100 slov za minutu [25].

K velkému rozvoji technologie rozpoznávání řeči došlo poté, co Geoffrey Hinton v roce 2006 znovu probudil zájem o výzkum hlubokých neuronových sítí (DNN). V roce 2011 Yu Dong a Deng Li ze společnosti Microsoft využili DNN pro kontinuální rozpoznávání řeči, čímž dosáhli výrazně nižší chybovosti. Toto zahájilo éru nového hybridního DNN-HMM přístupu, který kombinuje hluboké neuronové sítě a skryté Markovovy modely. Od roku 2014 začal získávat na popularitě end-to-end přístup, díky práci Navdeepa Jaitlyho a Alexe Gravesa z Google DeepMind, kteří představili model pro přímé převody vstupního akustického signálu na text. End-to-end systémy se v současnosti stávají stále efektivnějšími a přesnějšími [27].

### 3.3 Teoretický základ: HMM a DNN

#### 3.3.1 Skryté Markovovy modely (HMM)

Skryté Markovovy Modely (Hidden Markov Models), jsou statistické modely, které se používají k modelování sekvencí dat. Základem těchto modelů je Markovův řetězec,<sup>6</sup> jehož stavy nejsou přímo viditelné (jsou „skryté“). Skryté stavy jsou spojeny s pravděpodobnostní distribucí, která generuje pozorování, což jsou viditelné výstupy. Po vygenerování pozorování model podle přechodové pravděpodobnosti přechází do následujícího skrytého stavu. Hlavním cílem je odvodit nejpravděpodobnější sekvence skrytých stavů, jenž by mohly odpovídat vygenerovaným pozorováním. Model je definován maticí přechodových pravděpodobností mezi skrytými stavy, maticí emisních pravděpodobností<sup>7</sup> pro pozorování a pravděpodobnostmi počátečních stavů [28].

V kontextu rozpoznávání řeči je pro každé slovo nebo foném<sup>8</sup> potřeba vytvořit jeden model. Pro modelování slov lze tyto modely fonémů řetězit za sebe. Typickým přístupem k modelování fonémů je využití pětistavového levo-dopředního modelu HMM (viz obr. 13). Tento model má tři emitující stavy<sup>9</sup> ( $s_2, s_3, s_4$ ) a dva neemitující stavy ( $s_1, s_5$ ), z nichž první a poslední se obvykle využívají pro navazování na další modely. Přechodové pravděpodobnosti mezi stavy jsou na obr. 13 označeny jako „ $a_{ij}$ “ emisní pravděpodobnosti jsou označeny jako „ $b_{i(s)}$ “. V procesu rozpoznávání řečových sekvencí

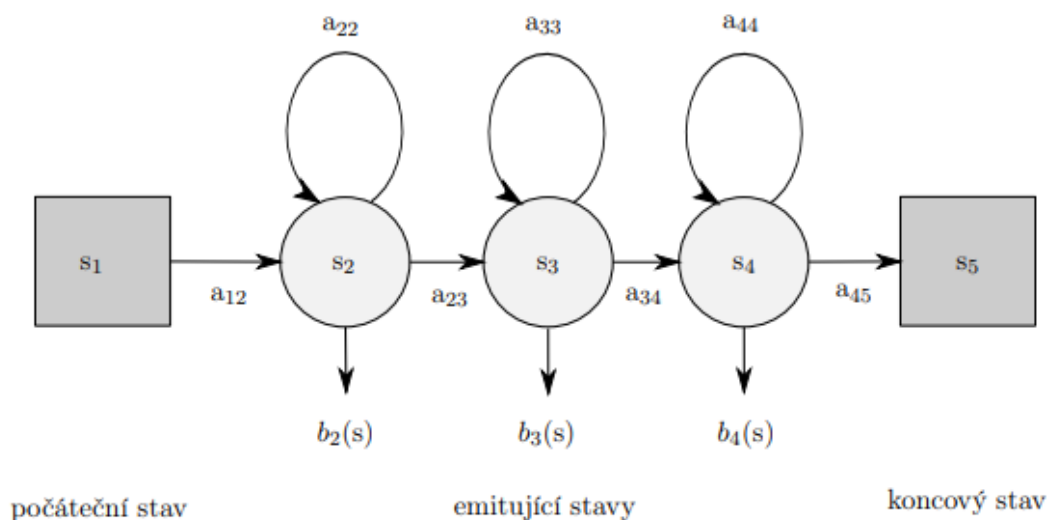
<sup>6</sup> Markovův řetězec je matematický model popisující systémy, které mění své stavy pouze na základě aktuálního stavu nikoli na předchozí historii [28].

<sup>7</sup> Matice emisních pravděpodobností určuje pravděpodobnost, s jakou se generují viditelné výstupy (emise) z každého ze skrytých stavů modelu. V kontextu rozpoznávání řeči by například ukazovala pravděpodobnost, s jakou se z daného fonému (skrytého stavu) vygeneruje určitý akustický signál (pozorovaný výstup) [28].

<sup>8</sup> Foném je základní jednotka zvuku v jazyce, která odlišuje význam slov.

<sup>9</sup> Emitující stav je ten, který vytváří pozorování, jinak řečeno viditelné výstupy modelu [28].

je cílem nalézt model, který s největší pravděpodobností generuje toto slovo nebo foném. Tuto pravděpodobnost lze vypočítat pomocí *Viterbiho algoritmu*<sup>10</sup> [29].



Obr. 13: Pětistavový levo-pravý HMM model [29]

### 3.3.2 Hluboké neuronové sítě (DNN)

Umělá neuronová síť představuje model strojového učení, který se pokouší napodobit způsob, jakým lidský mozek zpracovává informace. Inspirace pro tento model pochází ze struktury a funkce biologických neuronů v mozku. Základem umělé neuronové sítě je síť propojených uzlů neboli „umělých neuronů“ uspořádaných do několika vrstev, které zahrnují vstupní vrstvu, jednu nebo více skrytých vrstev a výstupní vrstvu. Vstupní vrstva přijímá data, skryté vrstvy tyto data zpracovávají a výstupní vrstva generuje konečný výsledek. Informace jsou mezi vrstvami přenášeny pomocí vážených spojení, jež určují, jak budou informace přenášeny. Dalším klíčovým prvkem každého neuronu ve vrstvách je aktivační funkce, která určuje, zda a jak silně bude neuron aktivován, a tím ovlivňuje přenos signálu do další vrstvy. Použití těchto funkcí vnáší do modelu nelinearitu, což umožňuje jeho učení složitých vzorců. Příklady těchto funkcí mohou být například sigmoid, hyperbolický tangent nebo ReLU.

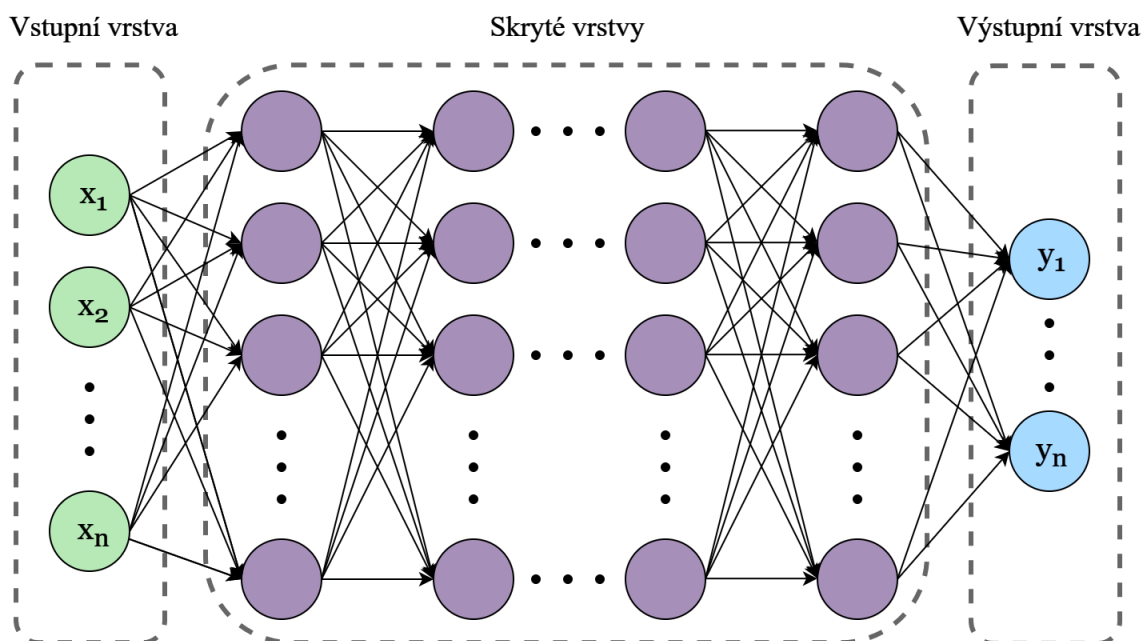
Trénink umělé neuronové sítě spočívá v optimalizaci vah spojení mezi neurony takovým způsobem, aby výstup sítě odpovídal co nejlépe očekávaným výstupům z trénovacích dat. Toto se obvykle provádí pomocí metody zvané *zpětné šíření chyb*, která využívá techniku gradientního sestupu k efektivnímu upravování vah tak, aby byla chyba v predikcích co nejmenší.

Základní umělé neuronové sítě mají obvykle jednu nebo dvě skryté vrstvy. Pokud však mají velké množství skrytých vrstev, často až desítky či stovky, lze je označit jako tzv. *hluboké neuronové sítě* (obr. 14). Hluboké neuronové sítě mají díky velkému

<sup>10</sup> Viterbiho algoritmus se používá k nalezení nejpravděpodobnější sekvence skrytých stavů, které mohly vygenerovat danou sekvenci pozorování [28].

množství skrytých vrstev schopnost učit se vysoce abstraktní vzorce v datech. Tato schopnost je činí vhodné právě například pro rozpoznávání řeči, nebo obrazu a další pokročilé aplikace. Na druhou stranu vyžadují větší množství trénovacích dat a složitější strategii pro zabránění přeučení.

Mezi základní typy hlubokých neuronových sítí patří konvoluční neuronové sítě (CNN), které jsou ideální pro zpracování obrazu díky své schopnosti z něho extrahovat rysy. Dalším typem jsou rekurentní neuronové sítě (RNN) vhodné obzvláště pro sekvenční úlohy, jako je například právě rozpoznávání řeči nebo zpracování textu. Díky interním smyčkám dokážou zachovávat informace z předchozích kroků během zpracování dat, což v podstatě simuluje „paměť“. S trénováním RNN se mohou pojít problémy, jako například zmizelý nebo explodující gradient. Tato skutečnost vedla k návrhu jejich pokročilých variant LSTM (Long Short-Term Memory) a GRU (Gated Recurrent Units), které jsou schopné uchovávat informace po delší dobu bez výrazné ztráty výkonu [30].



Obr. 14: Schéma hluboké neuronové sítě

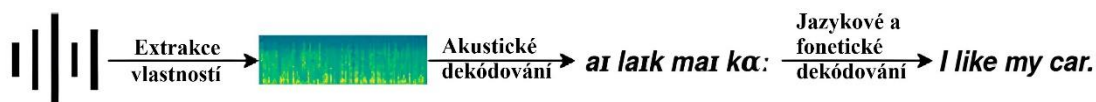
### 3.4 Metody rozpoznávání řeči

V této části jsou blíže popsány dva hlavní přístupy, a to tradiční systémy a moderní end-to-end systémy.

#### 3.4.1 Tradiční systémy

Tradiční (modulární) systémy pro rozpoznávání řeči jsou složeny z několika komponent a modelů, které jsou trénovány odděleně a následně spojeny do jednoho celku (obr. 15). Obvykle se skládají z akustického, fonetického a jazykového modelu [31].

- **Akustický model** – je klíčová komponenta jejíž úlohou je převod zvukového signálu na fonémy. Existují dva hlavní přístupy k jeho konstrukci: jeden využívá HMM ve spojení s modelem směsi Gaussových rozdělání (HMM-GMM), zatímco druhý kombinuje HMM s hlubokými neuronovými sítěmi (HMM-DNN). V současnosti je přístup HMM-DNN používanější z důvodu vyšší přesnosti a efektivity.
- **Fonetický model** – nazývaný také *slovník*, mapuje sekvence fonémů na slova nebo jiné jazykové jednotky a umožňuje systému porozumět různým dialektům a akcentům. Většinou bývá sestaven profesionálními lingvisty.<sup>11</sup>
- **Jazykový model** – predikuje, která slova nebo fráze po sobě nejpravděpodobněji následují. Tímto pomáhá sestavit věty gramaticky správně a zlepšuje přesnost. Obvykle je založen na statistických modelech, jako jsou N-gramy nebo na modelech rekurentních neuronových sítí [31; 32].



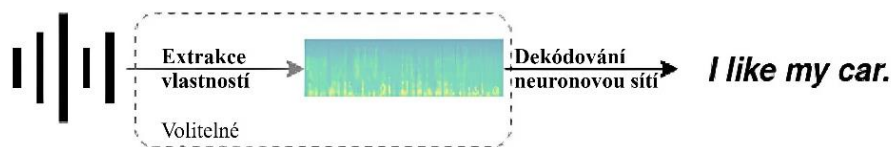
Obr. 15: Princip funkce tradičních modulárních systémů (popisky přeloženy autorem) [31]

#### 3.4.2 End-to-end systémy

End-to-end systémy představují moderní přístup značně odlišný od tradičních systémů. Tyto systémy jsou navrženy tak, aby převáděly vstupní audio signál přímo na textový výstup, čímž eliminují potřebu několika různých komponent typických pro tradiční přístup. Na rozdíl od tradičních systémů, které jsou složeny ze separátních modelů, end-to-end systémy integrují tyto modely do jediného komplexního modelu neuronové sítě. Tento unifikovaný model přímo spojuje vstup s výstupem (obr. 16). Trénink je prováděn na rozsáhlých datasetech, jenž obsahují páry nahrávek a odpovídajících textových přepisů [31].

<sup>11</sup> Lingvisté jsou odborníci studující jazyk a jeho aspekty, jako je fonetika, gramatika a další.





Obr. 16: Princip funkce end-to-end systémů (popisky přeloženy autorem) [31]

### 3.4.3 Porovnání tradičních a end-to-end systémů

Hlavní předností end-to-end systémů je jejich kompaktnost, která je výsledkem trénování a optimalizace celého systému jako jednotného celku. Na rozdíl od tradičních systémů, kde se optimalizuje každý model zvlášť, což často nevede k dosažení globálního optima celého systému. Díky kompaktnosti a absenci mezikroků mohou end-to-end systémy nabídnout i rychlejší zpracování a nižší latenci při rozpoznávání řeči [27].

Přes uvedené přednosti mají oproti tradičním systémům i několik nevýhod. Pro dosažení vynikajících výsledků vyžadují end-to-end systémy až desetinásobně či stonásobně více dat pro trénink, než je tomu u tradičních přístupů. To souvisí i s delší dobou trénování, která se pro rozsáhlé datasety o rozsahu tisíců hodin může v závislosti na dostupném výpočetním výkonu a velikosti modelu pohybovat od jednoho týdne až po několik měsíců. Navíc, pokud end-to-end systém produkuje chybné výsledky, je obtížné tyto chyby diagnostikovat a opravit, protože vzhledem k jeho „black-box“ charakteristice není lehké určit, která část je za chyby zodpovědná [33].



## 4 PRAKTICKÁ ČÁST

### 4.1 Výběr vhodného nástroje pro rozpoznávání řeči

Prvním důležitým krokem v procesu návrhu byl výběr adekvátního nástroje pro rozpoznání řeči, který by vyhovoval specifickým kritériím. Hledaný nástroj měl být buď zdarma, nebo komerční s možností získání bezplatného přístupu pro účely této bakalářské práce. Dalším důležitým kritériem byla také podpora českého jazyka, jelikož hlasové ovládání mělo probíhat v tomto jazyce. Po pečlivém průzkumu byly jako potenciální kandidáti splňující tyto kritéria vybrány nástroje Whisper od společnosti OpenAI a Google Speech Recognition API prostřednictvím knihovny SpeechRecognition. Kromě těchto dvou bezplatných nástrojů byly zvažovány i ty placené, jako například Google Cloud Speech-To-Text API, Amazon Transcribe a IBM Watson Speech to Text.

S žádostí o poskytnutí bezplatného přístupu k těmto službám byly kontaktovány společnosti Google, Amazon i IBM. Na žádost zareagovala společnost IBM, jejíž český zástupce poskytl instrukce pro přístup k jejich API bez nutnosti zadání platebních údajů z kreditní karty. Tato možnost však byla s jistými limitacemi, bylo k dispozici maximálně 500 minut rozpoznávání řeči měsíčně nebo alternativní limit 10 000 znaků měsíčně. Ostatní společnosti bohužel neodpověděly. Mezi zvažované možnosti tedy patřily Whisper, Google Speech Recognition API a IBM Watson Speech to Text. Následně bylo provedeno srovnávací testování vybraných nástrojů, aby bylo možné určit ten nejvhodnější pro účely hlasového ovládání kobota. V následujících podkapitolách jsou tyto jednotlivé nástroje představeny podrobněji a je popsán proces výběru toho nejvhodnějšího.

#### 4.1.1 Whisper

Whisper je model strojového učení od společnosti OpenAI, jenž byl vyvinut pro převod řeči na text. Jedná se o architekturu s end-to-end přístupem. Byl vydán jako open-source software v září roku 2022. Je schopen přepisovat řeč v mnoha jazycích, včetně češtiny. Současně dokáže překládat mluvené slovo z jednoho jazyka do jiného a může být použit i k identifikaci jazyka. Při jeho trénování bylo použito 680 000 hodin vícejazyčných zvukových supervizovaných dat z internetu, přičemž 117 000 hodin tvořila neanglická zvuková data. Výsledná přesnost se u různých jazyků liší, u těch, které nejsou dostatečně zastoupeny v trénovacích datech je chybovost přepisu slov vyšší. OpenAI uvedlo, že použití takto velkého a pestrého datasetu vede k dosažení lepšího rozpoznávání přízvuků, technického jazyka a vyšší odolnosti vůči rušivým vlivům šumu v okolí. Jeho velkou výhodou je, že ho lze použít bez připojení k internetu a je možné ho nasadit i na svých vlastních serverech, což se hodí především z hlediska ochrany dat. Naopak nevýhodou mohou být vysoké nároky na výpočetní výkon, při používání velkých verzí modelu je doporučeno použít výkonný grafický procesor [34; 35].

## Dostupné modely

Whisper nabízí pět velikostí modelu s různými kompromisy mezi rychlostí a přesností. Relativní rychlost je v tabulce 2 vyjádřena v porovnání s velkým modelem. Například střední model by měl být přibližně dvakrát rychlejší než velký model, avšak s očekávaným mírným poklesem přesnosti. Pro anglicky orientované aplikace jsou k dispozici modely s koncovkou „.en“, které obvykle poskytují větší přesnost než jejich univerzální protějšky. Tento rozdíl je zvláště patrný u drobného a základního modelu [35].

Tab. 2: Dostupné modely Whisperu [35]

Velikost	Parametry	Pouze anglický model	Vícejazyčný model	Požadovaná VRAM	Relativní rychlost
Drobný	39M	tiny.en	tiny	~1GB	~32×
Základní	74M	base.en	base	~1GB	~16×
Malý	244M	small.en	small	~2GB	~6×
Střední	769M	medium.en	medium	~5GB	~2×
Velký	1550M		large	~10GB	1×

### 4.1.2 IBM Watson Speech to Text

IBM Watson Speech to Text je komerční cloudová služba nabízející API pro převod mluvené řeči na text. Tato služba využívá technologii rozpoznávání řeči od společnosti IBM. Umožňuje přepis řeči na text ve více jazycích a z různých zvukových formátů pomocí strojového učení založeného na hlubokém učení. Konkrétní typ architektury se bohužel ve veřejných zdrojích nepodařilo najít. Slovník základního modelu obsahuje širokou škálu slov pro běžné každodenní konverzace, ale mohou v něm chybět některé odborné termíny ve specifických odvětvích jako je strojírenství, IT, medicína, právo a další. Z toho důvodu IBM poskytuje možnost přizpůsobení jazykového modelu prostřednictvím speciálního rozhraní, které umožňuje rozšíření základního slovníku o odbornou terminologii. Avšak pro využití tohoto rozhraní je nezbytné disponovat aktivním cenovým plánem Plus, Standard nebo Premium. V rámci této práce je k dispozici pouze plán Lite. Služba nabízí tři přístupové možnosti: WebSocket, synchronní a asynchronní HTTP. Rozhraní WebSocket nabízí plně duplexní spojení, umožňuje asynchronní komunikaci s možností odesílání více požadavků a přijímání výsledků přes jediné spojení. Synchronní HTTP rozhraní používá základní přístup k přepisu zvuku pomocí blokujících požadavků, což je ideální pro jednorázovou transkripci. Naopak asynchronní HTTP rozhraní podporuje neblokující přístup a je vhodné pro zpracování většího objemu dat [36].

## Dostupné modely

S účinností od 31. července 2023 služba IBM Watson Speech to Text podporuje pouze modely nové generace, které ve svých názvech obsahují termíny Telephony a Multimedia (viz tab. 3 a 4). Modely Telephony jsou navrženy primárně pro zvuk přenášený prostřednictvím mobilních telefonů s minimální vzorkovací frekvencí 8 kHz. Modely Multimedia jsou určeny pro zvuk získaný ze zdrojů s vyšší vzorkovací frekvencí, například z videa, s minimální vzorkovací frekvencí 16 kHz. Většina modelů nové generace rovněž nabízí možnost nízké latence, což umožňuje rychlejší získávání výsledků. Je však třeba brát v úvahu, že nízká latence může ovlivnit přesnost přepisu. Název modelu je složen z jazyka, ve kterém je zvuk namluven, a rychlosti, s jakou je vzorkován. V případě českého modelu je momentálně dostupná pouze varianta Telephony [36].

Tab. 3: Dostupné „Telephony“ modely služby IBM Watson Speech to Text [36]

Jazyk	Název modelu
Arabština (moderní standard)	ar-SA_Telephony
Čínština (mandarínská)	zh-CN_Telephony
Čeština	cs-CZ_Telephony
Holandština (belgická)	nl-BE_Telephony
Holandština (nizozemská)	nl-NL_Telephony
Angličtina (australská)	en-AU_Telephony
Angličtina (indická)	en-IN_Telephony
Angličtina (Spojené království)	en-GB_Telephony
Angličtina (Spojené státy americké)	en-US_Telephony
Angličtina (medicínská)	en-WW_Medical_Telephony
Francouzština (kanadská)	fr-CA_Telephony
Francouzština (Francie)	fr-FR_Telephony
Němčina	de-DE_Telephony
Hindština (indická)	hi-IN_Telephony
Italština	it-IT_Telephony
Japonština	ja-JP_Telephony
Korejština	ko-KR_Telephony
Portugalština (brazilská)	pt-BR_Telephony
Španělština (kastilská)	es-ES_Telephony
Španělština (argentinská, čilská, kolumbijská, mexická, peruánská)	es-LA_Telephony
Švédština	sv-SE_Telephony

Tab. 4: Dostupné „Multimedia“ modely služby IBM Watson Speech to Text [36]

Jazyk	Název modelu
Holandština (nizozemská)	nl-NL_Multimedia
Angličtina (australská)	en-AU_Multimedia
Angličtina (Spojené království)	en-GB_Multimedia
Angličtina (Spojené státy)	en-US_Multimedia
Francouzština (kanadská)	fr-CA_Multimedia
Francouzština (Francie)	fr-FR_Multimedia
Němčina	de-DE_Multimedia
Italština	it-IT_Multimedia
Japonština	ja-JP_Multimedia
Korejština	ko-KR_Multimedia
Portugalský (brazilský)	pt-BR_Multimedia
Španělština (kastilská)	es-ES_Multimedia

#### 4.1.3 Google Speech Recognition prostřednictvím knihovny SpeechRecognition

Knihovna SpeechRecognition představuje univerzální rozhraní, jenž umožňuje vývojářům v Pythonu snadno implementovat funkcionalitu rozpoznávání řeči do jejich projektů. Slouží jako obal (wrapper) pro řadu populárních systémů rozpoznávání řeči. Mezi tyto systémy v současnosti patří například:

- CMU Sphinx
- Google Speech Recognition
- Google Cloud Speech API
- Wit.ai
- Microsoft Azure Speech
- Houndify API
- IBM Watson Speech to Text
- Snowboy Hotword Detection
- Vosk API
- OpenAI whisper
- Whisper API [37]

Systém Google Speech Recognition, který je v tomto seznamu, využívá pro převod řeči na text Google Web Speech API s klíčem zabudovaným přímo v knihovně SpeechRecognition. To znamená, že tuto službu lze využívat zdarma, bez registrace nebo vlastního klíče. Je nutné podotknout, že výchozí klíč zabudovaný v této knihovně je pouze pro testovací účely a společnost Google ho může kdykoliv zneplatnit, proto není vhodné tuto službu využívat ke komerčním účelům. Google Web Speech API podporuje mnoho jazyků, včetně češtiny [38]. Informace o přesných technických podrobnostech a architektuře Google Web Speech API nebyly ve veřejně dostupných dokumentacích nalezeny.

#### 4.1.4 Porovnání jednotlivých systémů

V rámci výběru nejvhodnějšího nástroje pro rozpoznávání řeči byly testovány české modely všech již zmíněných systémů. Je důležité zmínit, že cílem nebylo provést komplexní porovnání schopností rozpoznávání řeči v různých kontextech nebo všech dostupných funkcionalit, ale spíše zjednodušené porovnání, které odpovídalo specifickým potřebám v kontextu hlasového ovládání pro kobota. Mezi hlavní hodnocená kritéria patřila přesnost a rychlost transkripce.

Pro účely testování byl vytvořen seznam 20 testovacích příkazů, zahrnujících řadu potenciálních reálných pokynů pro kobota (viz tab. 5). Při jejich výběru byl kladen důraz na různorodost slov. Následně byly tyto příkazy pro každý systém namluveny dvěma osobami, mužem a ženou, pro větší rozmanitost hlasových vzorů na mikrofon Superlux E205U. Během testování byla měřena doba potřebná k převodu řeči na text, a to od momentu zahájení rozpoznávání až po vygenerování příslušné transkripce. Výsledky byly zaznamenány do dvou tabulek, pro mužský a ženský hlas. V těchto tabulkách níže bylo vždy specifikováno číslo příkazu, přesný transkript poskytnutý systémem, ve kterém jsou tučně vyznačena slova neshodující se s původními testovacími příkazy, a čas potřebný k převodu nahrávky příkazu na příslušný přepis.

Tab. 5: Testovací příkazy

Číslo	Příkaz
1	Stop.
2	Vrať se do domovské pozice.
3	Zpomal svůj pohyb.
4	Zrychli svůj pohyb na maximální rychlost.
5	Umísti tento objekt na stůl.
6	Proveď rotaci nástroje o 45 stupňů.
7	Otoč se vlevo o 90 stupňů.
8	Sniž se o 15 centimetrů.
9	Posuň se vpřed o 33 centimetrů.
10	Zvedni se o 20 centimetrů.
11	Nastav pracovní rychlost na 50 %.
12	Přesuň se do polohy označené číslem 5.
13	Převed' se do režimu jemné manipulace.
14	Aktivuj režim manuálního řízení.
15	Vyčisti pracovní prostor.
16	Spusť automatický režim skládání.
17	Nastav teplotu nástroje na 200 stupňů Celsia.
18	Změň pracovní nástroj.
19	Zastav se.
20	Nastav výchozí polohu.

Tab. 6: Výsledky rozpoznání hlasových příkazů mužského respondenta pomocí modelu Whisper

Číslo	Rozpoznaný text (Whisper)	Čas (s)
1	Stop.	17,02
2	Vrať se do domovské pozice.	19,04
3	<b>spomal</b> svůj pohyb.	19,53
4	<b>zrychlý</b> svůj pohyb na maximální rychlost.	17,92
5	Umísti tento objekt na stůl.	17,82
6	Proveď rotaci nástroje o 45°.	18,46
7	Otoč se vlevo o 90°.	16,95
8	<b>Sníž</b> se o 15 cm.	17,03
9	<b>Posuní</b> se vpřed o 33 cm.	16,11
10	Zvedni se o 20 cm.	17,92
11	Nastav pracovní rychlost na 50%.	15,95
12	<b>Přesuň</b> se do polohy označené číslem 5.	19,49
13	Převeď se do režimu jemné manipulace.	18,95
14	Aktivuj režim manuálního řízení.	19,21
15	Vyčisti pracovní prostor.	18,68
16	<b>Spustí</b> automatický režim skládání.	19,18
17	Nastav teplotu nástroje na 200°C.	20,11
18	Změň pracovní nástroj.	20,99
19	<b>Zastav</b> se.	20,12
20	Nastav výchozí polohu.	18,63

Tab. 7: Výsledky rozpoznání hlasových příkazů ženské respondentky pomocí modelu Whisper

Číslo	Rozpoznaný text (Whisper)	Čas (s)
1	Stop.	18,22
2	Vrať se do domovské pozice.	17,89
3	<b>spomal</b> svůj pohyb.	19,07
4	<b>Zrychlý</b> svůj pohyb na maximální rychlost.	20,27
5	Umísti tento objekt na stůl.	17,56
6	Proveď rotaci nástroje o 45°.	18,66
7	Otoč se vlevo o 90°.	19,02
8	<b>Sníž</b> se o 15 cm.	17,05
9	Posuň se vpřed o 33 cm.	18,56
10	Zvedni se o 20 cm.	18,22
11	Nastav pracovní rychlost na 50%.	18,52
12	Přesuň se do polohy označené číslem 5.	18,71
13	Převeď se do režimu jemné manipulace.	20,38
14	Aktivuj režim manuálního řízení.	16,35
15	Vyčisti pracovní prostor.	20,66
16	<b>Spud'</b> automatický režim skládání.	19,83
17	Nastav teplotu nástroje na 200°C.	19,14
18	Změň pracovní nástroj.	18,69
19	<b>Zastav</b> se.	20,12
20	Nastav výchozí polohu.	18,75



Tab. 8: Výsledky služby IBM Watson Speech to Text pro hlasové příkazy mužského respondenta

Číslo	Rozpoznaný text (Watson)	Čas (s)
1	stop	2,63
2	vrať se do domovské pozice	2,21
3	zpomal svůj pohyb	2,23
4	zrychli svůj pohyb na maximální rychlost	2,7
5	<b>umístí</b> tento <b>oběk</b> na stůl	2,57
6	proved' <b>rotoci</b> nástroje o čtyřicet pět <b>vstup</b>	3,45
7	otoč se vlevo o devadesát <b>tupilů</b>	2,85
8	<b>sníž</b> se o patnáct centimetrů	2,4
9	<b>osmý set před</b> o třicet tři centimetrů	2,09
10	<b>zvední</b> se o dvacet centimetrů	2,74
11	<b>na stav</b> pracovní rychlost na padesát procent	2,45
12	<b>přesuní</b> se do polohy označené číslem pět	2,32
13	převედ' se <b>doryžimu je mné</b> manipulace	2,41
14	<b>aktivuji režim</b> manuálního řízení	2,39
15	<b>vy čistí</b> pracovní prostor	2,95
16	<b>spustí</b> automatický režim skládání	2,28
17	<b>na stav té plotu</b> nástroje na dvě stě <b>vstupný ucelzia</b>	2,15
18	<b>změní</b> pracovní nástroj	2,44
19	zastav <b>te</b>	2,22
20	nastav výchozí polohu	2,31

Tab. 9: Výsledky služby IBM Watson Speech to Text pro hlasové příkazy ženské respondentky

Číslo	Rozpoznaný text (Watson)	Čas (s)
1	stop	2,24
2	vrať se do domovské pozice	2,36
3	zpomal svůj pohyb	2,26
4	zrychli svůj pohyb na maximální rychlost	2,94
5	<b>umístěn</b> tento <b>oběk</b> na stůl	2,35
6	proved' <b>rota ty</b> nástroje o čtyřicet pět <b>vstupnů</b>	2,08
7	otoč se vlevo o <b>devadesátu dnů</b>	2,72
8	<b>snížít</b> se o patnáct centimetrů	2,56
9	<b>osm set před</b> o třicet tři centimetrů	2,36
10	<b>zvední</b> se o dvacet centimetrů	2,48
11	<b>na stav</b> pracovní rychlost na padesát procent	2,39
12	<b>přesuně</b> se do polohy označené číslem pět	2,32
13	převעד' se <b>dorazimu je mne</b> manipulace	2,62
14	aktivuj <b>ražím</b> manuálního řízení	2,33
15	<b>ve čistě</b> pracovní prostor	2,56
16	<b>spust</b> automatický režim skládání	2,66
17	<b>na stav té plutu</b> nástroje na dvě stě <b>stupnů ucelzia</b>	2,74
18	<b>změní</b> pracovní nástroj	2,84
19	<b>v zastavce</b>	2,11
20	<b>no stav</b> výchozí polohu	2,23

Tab. 10: Výsledky služby Google Speech Recognition pro hlasové příkazy mužského respondenta

Číslo	Rozpoznaný text (Google)	Čas (s)
1	stop	2,14
2	Vrať se do domovské pozice	3,06
3	zpomal svůj pohyb	5,55
4	Zrychli svůj pohyb na maximální rychlost	2,8
5	<b>umístit</b> tento objekt na stůl	2,9
6	proved' rotaci nástroje o 45 stupňů	3,27
7	otoč se vlevo <b>od</b> 90 stupňů	3,88
8	sniž se o 15 cm	4,34
9	posuň se vpřed o 33 cm	2,14
10	Zvedni se o 20 cm	2,83
11	Nastav pracovní rychlost na 50%	3,93
12	Přesuň se do polohy označené číslem 5	2,97
13	převeď se do režimu jemné manipulace	2,01
14	Aktivuj režim manuálního řízení	2,36
15	vyčisti pracovní prostor	2,56
16	Spust' automatický režim skládání	4,12
17	Nastav teplotu nástroje na 200 stupňů Celsia	2,96
18	změň pracovní nástroj	2,72
19	zastav se	2,91
20	Nastav výchozí polohu	1,96

Tab. 11 Výsledky služby Google Speech Recognition pro hlasové příkazy ženské respondentky

Číslo	Rozpoznaný text (Google)	Čas (s)
1	stop	2,21
2	Vrať se do domovské pozice	3,52
3	zpomal svůj pohyb	2,27
4	Zrychli svůj pohyb na maximální rychlost	2,81
5	<b>umístit</b> tento objekt na stůl	3,11
6	proved' rotaci nástroje o 45 stupňů	3,21
7	otoč se vlevo <b>od</b> 90 stupňů	2,08
8	sniž se o 15 cm	2,84
9	posuň se vpřed o 33 cm	2,31
10	Zvedni se o 20 cm	4,95
11	Nastav pracovní rychlost na 50%	2,84
12	Přesuň se do polohy označené číslem 5	3,13
13	převeď se do režimu jemné manipulace	2,65
14	Aktivuj režim manuálního řízení	1,97
15	vyčisti pracovní prostor	2,59
16	Spust' automatický režim skládání	3,12
17	Nastav teplotu nástroje na 200 stupňů Celsia	2,82
18	změň pracovní nástroj	2,14
19	zastav se	1,61
20	Nastav výchozí polohu	1,94

## Výsledky porovnání

Pro prezentaci výsledků každého systému, byla data z tabulek pro mužský a ženský hlas sloučena do jedné souhrnné tabulky. Tato tabulka obsahuje průměrné hodnoty pro každý testovaný systém (tab. 12). Přesnost rozpoznávání je zde vyjádřena v procentech, jako poměr správně rozpoznávaných slov vůči celkovému počtu slov.

Tab. 12: Vyhodnocení testování

<b>Systémy rozpoznávání řeči</b>	<b>Přesnost (%)</b>	<b>Průměrný čas (s)</b>
Whisper	93,55	18,62
IBM Watson Speech To Text	72,04	2,47
Google Speech Recognition	97,85	2,89

Z hlediska přesnosti dosáhl nejhorších výsledků systém IBM Watson Speech to Text. Jeho přesnost by se však mohla značně zlepšit s využitím možnosti přizpůsobení jazykového modelu v případě cenového plánu Plus a vyššího. Na druhou stranu byla rychlost tohoto systému ze všech testovaných nejvyšší. Výrazně vyšší přesnosti dosáhl systém Whisper, využívající vícejazyčný model střední velikosti. Nicméně průměrný čas potřebný k převodu hlasového příkazu na text byl zhruba 7× delší ve srovnání se zbývajícími dvěma testovanými systémy. Tento rozdíl byl způsoben pravděpodobně tím, že testování probíhalo off-line na notebooku vybaveným procesorem Intel(R) Core (TM) i7-11370H a 16 GB RAM, bez využití grafické karty. V případě použití výkonnějšího počítače by byl čas potřebný k převodu kratší. Nejlepších výsledků z pohledu přesnosti dosáhl systém Google Speech Recognition API. Jeho rychlost byla jen zanedbatelně nižší oproti systému IBM Watson Speech to Text, navíc ho lze využívat zdarma. Z těchto důvodů bude právě tento systém využit v programu pro hlasové ovládání kobota.

## 4.2 Návrh programu pro hlasové ovládání kolaborativního robota

### 4.2.1 Požadavky na program

Cílem programu mělo být umožnit obsluhu efektivně a bezpečně ovládat kolaborativního robota UR3 prostřednictvím hlasových příkazů. Princip ovládání měl být jednoduchý, s předem nadefinovanými činnostmi spouštěnými jednotlivými hlasovými příkazy. Nejdůležitějším požadavkem na program byla bezpečnostní opatření. Vzhledem k možné přítomnosti mnoha osob na pracovišti hrozí, že by v průběhu běžné konverzace mohl kdokoliv neúmyslně vyslovit příkaz pro kobota a nechtěně ho rozpohybovat v nežádoucí situaci. To by mohlo vést k poškození majetku nebo ohrožení obsluhy. Proto bylo vyžadováno, aby se hlasové ovládání aktivovalo až po vyřknutí specifického aktivačního příkazu, po jehož zaregistrování by program zároveň poskytl zpětnou vazbu a upozornil tak obsluhu, že je připraven přijmout příkaz. Dalším aspektem bylo, že kobota může být implementován v hlučném prostředí, kde může vznikat šum. Z toho důvodu byla dalším

požadavkem schopnost programu se tomuto hluku dokázat alespoň částečně přizpůsobit. Program musel být přehledný, snadno rozšiřitelný a muselo v něm být možné snadno měnit různé parametry i bez jeho hlubší znalosti.

#### 4.2.2 Volba nástrojů

Program byl napsán v programovacím jazyce *Python*. Jedná se o vysokoúrovňový interpretovaný programovací jazyk, který je velmi oblíbený především díky bohatému ekosystému knihoven a frameworků a jednoduché syntaxi. Python podporuje několik programovacích paradigmat, včetně objektově orientovaného, procedurálního a funkcionálního programování. Jedná se také o multiplatformní programovací jazyk, to znamená, že programy napsané v tomto jazyce mohou být spouštěny na různých operačních systémech bez nutnosti modifikací [39].

Pro převod hlasových příkazů na text byl po předchozím testování zvolen systém Google Speech Recognition skrze knihovnu *SpeechRecognition*. Tento systém je popsán podrobněji v kapitole 4.1.3. Knihovna *SpeechRecognition* také poskytla řešení pro adaptaci systému na hlučné prostředí. Tímto řešením bylo využití její metody *adjust\_for\_ambient\_noise* ze třídy *recognizer*, která dokáže automaticky nastavit citlivost mikrofону s cílem minimalizovat vliv okolního hluku na proces rozpoznávání řeči. Princip její funkce spočívá v krátkodobém naslouchání okolnímu prostředí bez toho, aniž aby se snažila rozpoznat nějaký hlas. Během naslouchání analyzuje úroveň hluku a nastaví citlivost tak, aby rozpoznávač mohl efektivně rozlišit hlas a šum v pozadí [38]. Následkem využití této funkce je větší zpoždění rozpoznávání řeči, proto je vhodné zvážit, kdy ji má smysl využít.

Pro ovládání kobota byla využita knihovna *ur\_rtde*, která umožňuje efektivní komunikaci s koboty Universal Robots prostřednictvím protokolu Real-Time Data Exchange (RTDE). Je kompatibilní s operačními systémy Windows, Linux i macOS a podporuje jazyky Python a C++ [40].

Pro přehrávání zvuku při aktivaci byla zvolena knihovna *pygame*. Dále byly využity dvě standardní knihovny jazyka Python, *Math* pro provádění matematických operací a *Time* pro manipulaci s časem.

#### 4.2.3 Funkce a architektura navrženého programu

##### Popis funkce programu

Před spuštěním samotného programu je nutné nakonfigurovat klíčové parametry v konfiguračním souboru. Tato konfigurace zahrnuje definici IP adresy robota, nastavení aktivačního slova, které je v tomto případě „*robote*“, a určení jazyka pro rozpoznávání řeči, zde češtiny (cs-CZ). Dále je třeba specifikovat cestu k zvukovému souboru, jenž se použije jako signál pro uživatele o připravenosti programu přijímat hlasové příkazy. Nastavení musí zároveň obsahovat volbu (True/False), zda se má využít dynamické přizpůsobení citlivosti mikrofónu k úrovni hluku v prostředí a definovaný počet pokusů, které má uživatel k dispozici pro zadání platného hlasového příkazu.

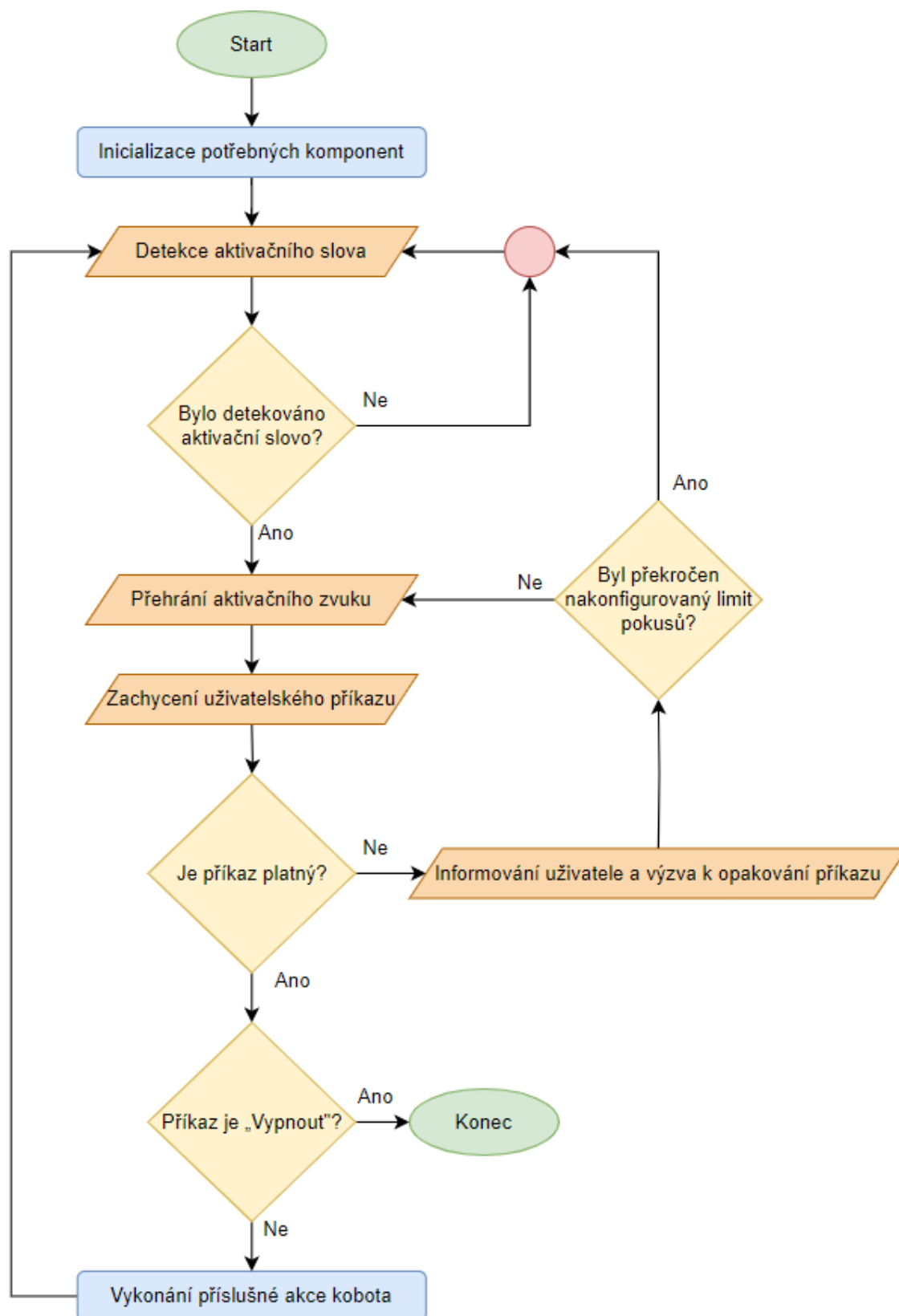
Po úspěšné konfiguraci a spuštění programu se nejprve inicializují potřebné komponenty pro zpracování hlasových příkazů a ovládání kobota. Tento krok zahrnuje navázání spojení s kobotem a přípravu audio systému pro přehrávání aktivačního zvuku.

Následně se spustí hlavní smyčka, přičemž program přejde do režimu, kdy aktivně naslouchá a monitoruje zvukové prostředí a snaží se zachytit aktivační slovo. Když se programu podaří detekovat aktivační slovo, signalizuje, že poslouchá a je připraven přijímat příkazy přehráním aktivačního zvuku. V ten moment má uživatel možnost vyslovit hlasový příkaz, který program prostřednictvím zvoleného systému rozpoznávání řeči převede na textovou formu. Tu dále zpracovává k vykonání příslušných akcí kobota, jakou jsou zde pro demonstraci pohyby do určených pozic, manipulace s objekty nebo provádění jiných specifických úloh.

V případě přeroku, nebo vyslovení v systému nedefinovaného příkazu, je uživatel informován a vyzván k jeho zopakování. Díky konfiguračnímu nastavení počtu pokusů pro zadání platného příkazu má uživatel několik šancí k opravě, aniž by bylo nutné znovu pronášet aktivační slovo. Po úspěšném zpracování příkazu nebo vyčerpání povolených pokusů se program automaticky vrátí do stavu poslechu, a je připraven opět reagovat na aktivační slovo a další pokyny. Tento cyklus se opakuje, čímž se zajišťuje plynulá a efektivní interakce mezi uživatelem a robotem. Celý princip funkce programu je znázorněn pomocí vývojového diagramu na obr. 17.

Pro demonstraci funkčnosti programu hlasového ovládání bylo v tomto případě pro kolaborativního robota definováno osm specifických příkazů:

1. *Domovská pozice* – Přesune se do domovské pozice.
2. *Běž do bodu A* – Přesune se do nedefinovaného bodu A.
3. *Přesuň se do bodu B* – Přesune se do nedefinovaného bodu B.
4. *Běž do pozice 3* – Přesune se do nedefinované pozice 3.
5. *Polož předmět* – Naznačí položení předmětu.
6. *Zvedni předmět* – Naznačí zvednutí předmětu.
7. *Pozdrav* – Zamává.
8. *Otoč se o 180°* – Otočí se o 180° vůči jeho současné pozici.
9. *Výpnout* – Dojde k ukončení programu.



Obr. 17: Vývojový diagram znázorňující princip funkce programu

## Architektura programu

Pro lepší přehlednost, rozšiřitelnost a flexibilitu byla zvolena modulární architektura. Program je rozdělen do čtyř komponent:

- **Konfigurační soubor** – Slouží jako centrální místo pro správu všech klíčových nastavení, která ovlivňují funkčnost a chování programu. Díky tomu lze všechny parametry snadno upravovat bez nutnosti zasahovat do hlavního kódu.
- **Modul pro zpracovávání hlasových příkazů** – Zodpovídá za detekci aktivačního příkazu a zpracování hlasových příkazů.
- **Modul pro ovládání robota** – Stará se o komunikaci s robotem, řídí jeho pohyby a další operace založené na instrukcích získaných ze zpracování hlasových příkazů.
- **Hlavní modul** – Řídí celkový běh programu a koordinuje interakce mezi jednotlivými komponentami. Obsahuje hlavní smyčku programu, která čeká na aktivaci a zpracovává příkazy.

### 4.3 Testování funkčnosti řešení pomocí simulace a v laboratoři

V této etapě bylo nutné ověřit správnou funkčnost celého systému. Testování navrženého programu probíhalo nejprve v simulovaném prostředí s virtuálním kobotem. Tento přístup byl zvolen ze dvou důvodů. Prvním důvodem byla možnost odladění pohybových sekvencí kobota tak, aby se předešlo nežádoucím kolizím. Druhým důvodem bylo, že poskytoval možnost pohodlného testování z domova, čímž se eliminovala potřeba časté fyzické přítomnosti v laboratoři. Po úspěšném ověření funkčnosti programu v simulaci následovalo testování na fyzickém robotovi v robotické laboratoři na Ústavu automatizace a informatiky.

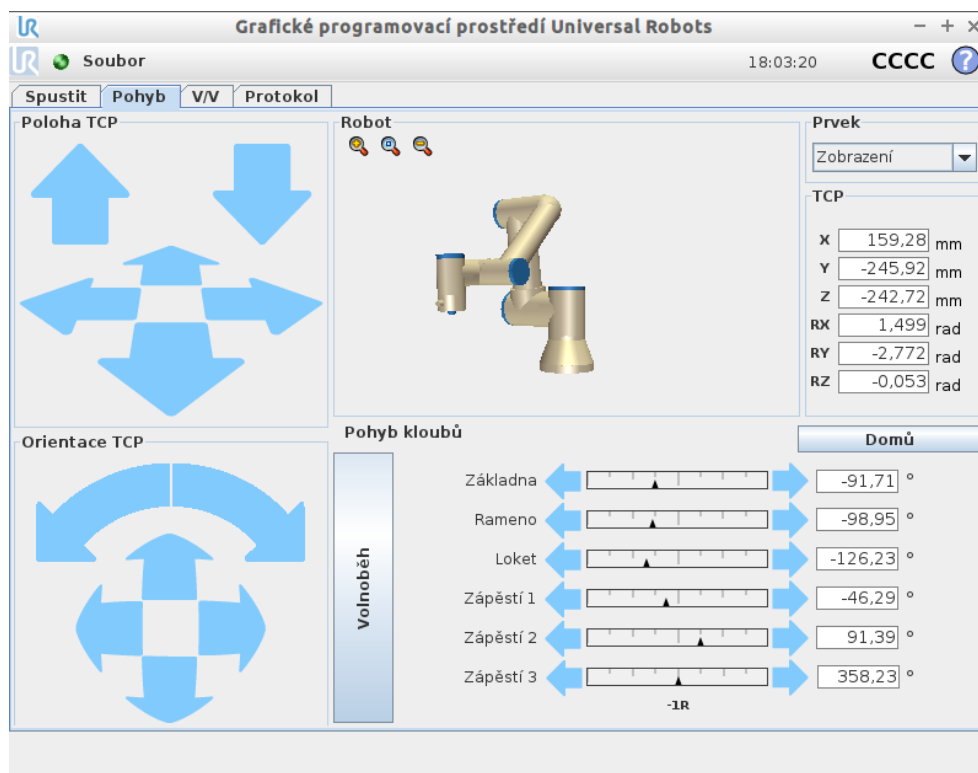
V obou případech byl program spouštěn na běžném notebooku s operačním systémem Windows 11 Pro ve virtuálním stroji s operačním systémem Ubuntu, což je distribuce založená na Linuxu. K záznamu zvuku byl použit stolní mikrofon Superlux E205U.

#### Testování v simulaci

Pro simulaci kobota byl využit off-line simulační software URSim 3.15.8 CB-SERIES. Tento software, vyvinutý pro operační systém Linux, umožňuje uživatelům vytvářet, testovat a simulovat programy pro kolaborativní roboty od společnosti Universal Robots off-line ve virtuálním prostředí. Aby bylo možné spustit URSim na operačním systému Windows, bylo nezbytné použít virtualizační software VMware Workstation Pro, který umožňuje vytvoření virtuálního stroje s operačním systémem Linux, na němž bylo možné URSim bez problému spustit.

Po úspěšném spuštění URSim se na domovské stránce zobrazí výběr ze tří simulátorů, každý pro jiný typ kolaborativního robota ze série CB, a to pro UR3, UR5 a UR10. V tomto případě byl vybrán simulátor pro kobota UR3. Poté se načte uživatelské

rozhraní Polyscope (viz obr. 18), ve kterém lze pracovat stejně jako při on-line programování pomocí Teach Pendantsu. Dalším krokem bylo zadat IP adresu kobota ze simulace do konfiguračního souboru v programu pro hlasové ovládání, čímž se program propojil se simulací a dokázal s kobotem pohybovat. Celý systém fungoval bez problémů a splňoval veškerá očekávání. Bylo nutné pouze upravit souřadnice některých bodů pro demonstrační úlohy.

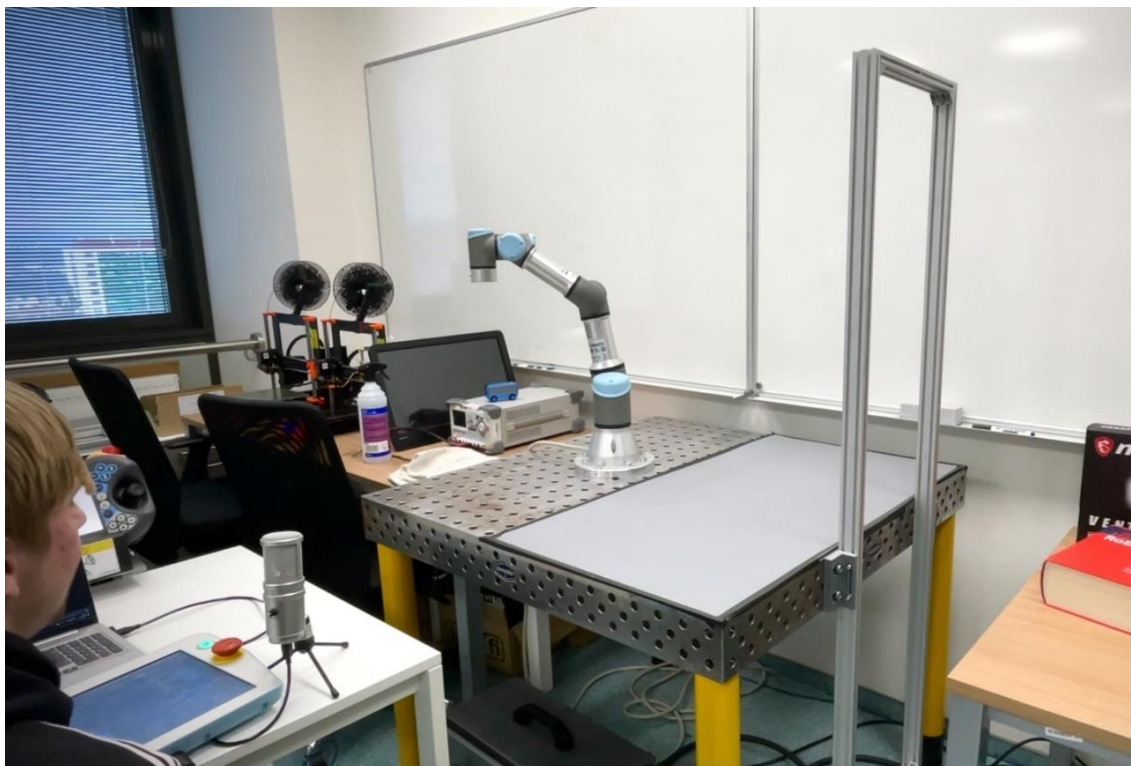


Obr. 18: Uživatelské prostředí Polyscope v simulátoru URSim 3.15.8 pro sérii CB

### Testování v laboratoři

Po úpravách souřadnic několika bodů v programu v rámci simulačního prostředí a ověření, že nehrozí riziko kolize, byl program otestován v laboratoři na reálném kolaborativním robotovi UR3 (viz obr. 19). Příprava na testování zahrnovala propojení řídicí jednotky kobota s místním routerem pomocí ethernetového kabelu, přičemž notebook byl připojen ke stejné síti prostřednictvím Wi-Fi. Následně byl kobot spuštěn a na ovládacím panelu byla zjištěna jeho IP adresa, která byla následně zapsána do konfiguračního souboru programu pro hlasové ovládání. Po navázání komunikace mezi programem a kobotem bylo zahájeno testování. Všechny operace proběhly bez problémů a přesně v souladu s výsledky dosaženými v simulaci. Testování bylo natočeno a záznam je obsažen v příloze.





Obr. 19: Testování v robotické laboratoři na ÚAI



## 5 DISKUZE

V této bakalářské práci byl navržen a implementován systém hlasového ovládání pro kolaborativního robota UR3 od společnosti Universal Robots. Výsledky testování ukázaly, že tento systém dokáže efektivně rozpoznávat hlasové příkazy a následně provádět odpovídající úkony. Přestože je systém funkční, nepochybně existuje prostor pro další vylepšení v budoucnosti.

Vývoj systému hlasového ovládání byl omezen na využití neplacené služby pro rozpoznávání řeči. Implementací komerční služby pro rozpoznávání řeči, například Google Cloud Speech-To-Text, by se pravděpodobně zlepšila přesnost a rychlost transkripce pro český jazyk. V rámci této práce byl proveden pokus o vývoj vlastního modelu neuronové sítě pro rozpoznávání řeči v českém jazyce pomocí TensorFlow. Záměrem bylo vytvořit model, který by sloužil jako alternativa ke službám pro rozpoznávání řeči využívaných v této práci. Navrhovaný model obsahoval dvě konvoluční vrstvy (CNN) s MaxPoolingem, za nimiž následovaly dvě vrstvy LSTM pro sekvenční zpracování dat, a dále dvě plně propojené vrstvy (Dense), přičemž druhá z nich působila jako výstupní vrstva s aktivací softmax. Jako zdroj trénovacích dat byl vybrán dataset CommonVoice, poskytovaný organizací Mozilla. Tento dataset obsahoval přibližně 70 hodin nahrávek v českém jazyce o celkové velikosti přibližně 5 GB ve formátu MP3. Nicméně, model nedosáhl požadované přesnosti, pravděpodobně v důsledku omezeného množství trénovacích dat pro český jazyk. Z toho důvodu nebyl v této práci dále využit. Do budoucna by však bylo vhodné zpracovat rozsáhlejší dataset a pokusit se o vývoj takového modelu.

Dalším vylepšením by mohla být integrace pokročilých technologií zpracování přirozeného jazyka, jako je model ChatGPT od OpenAI. Tato technologie by uživatelům umožnila zadávat příkazy svými slovy, přičemž systém by byl schopen rozpoznat kontext a provést odpovídající akce. To by vedlo k výrazně intuitivnějšímu hlasovému ovládání.

Doporučuje se také vytvoření intuitivního a přívětivého uživatelského rozhraní, které by zjednodušilo práci se systémem. Integrace vizuálních prvků a zpětné vazby pro uživatele by mohla významně přispět ke zlepšení celkové uživatelské zkušenosti.

Na závěr by bylo vhodné tento systém přizpůsobit pro reálnou aplikaci v praxi a otestovat jeho účinnost a přínosnost. Budoucí výzkum by se měl zaměřit na jeho vylepšení a následnou implementaci a testování v různých odvětvích, například ve výrobě, logistice, zdravotnictví nebo osobní asistenci, a vyhodnotit, do jaké míry je účinné a přínosné hlasové ovládání kolaborativních robotů v těchto oblastech.



## 6 ZÁVĚR

Cílem této bakalářské práce bylo seznámit se s oblastí kolaborativní robotiky a technologií rozpoznávání řeči, a následně navrhnout a implementovat systém hlasového ovládání pro kolaborativního robota UR3 od společnosti Universal Robots.

V teoretické části byla čtenářům představena kolaborativní robotika, její historie, aplikace v různých odvětvích a metody programování kolaborativních robotů. Rovněž byla představena společnost Universal Robots a detailně popsán model UR3 CB série. Dále byla rozebrána technologie rozpoznávání řeči, včetně jejího současného využití, historického vývoje a metod.

V praktické části byly představeny a porovnány existující systémy rozpoznávání řeči, které splňovaly zadané požadavky, a následně byl vybrán nejvhodnější z nich. Na základě testování byl jako nejvhodnější vybrán systém Google Speech Recognition, implementovaný prostřednictvím knihovny SpeechRecognition. Poté byl vytvořen program v programovacím jazyce Python pro hlasové ovládání kolaborativního robota, jehož návrh, funkce a architektura byly podrobně popsány. Výsledný systém byl úspěšně otestován nejprve v simulaci a poté i v robotické laboratoři na Ústavu automatizace a informatiky.

Výsledky testování ukázaly, že navržený systém hlasového ovládání je schopen spolehlivě rozpoznávat a provádět zadané příkazy. Videodokumentaci z testování lze nalézt v příloze.



## SEZNAM POUŽITÉ LITERATURY

- [1] VOJÁČEK, Antonín. Problematika bezpečnosti kolaborativních robotů - ISO/TS 15066. Online. Automatizace.HW.cz. 2019. Dostupné z: <https://automatizace.hw.cz/problematika-bezpecnosti-kolaborativnich-robotu-isots-15066.html>. [cit. 2024-04-06].
- [2] KARLÍK, Tomáš. Před sto lety svět poprvé uslyšel slovo „robot“. Karel Čapek inspiroval ke vzniku strojů, které vládou dnešním továrnám. Online. ČT24. 2021. Dostupné z: <https://ct24.ceskatelevize.cz/clanek/veda/pred-sto-lety-svet-poprve-uslysel-slovo-robot-karel-capek-inspiroval-ke-vzniku-stroju-ktere-vladnou-39198>. [cit. 2024-04-06].
- [3] UNIVERSAL ROBOTS. Inventor of the Robot Arm and Its Continued Development. Online. Universal Robots. 2020. Dostupné z: <https://www.universal-robots.com/blog/inventor-of-the-robot-arm-and-its-continued-development/>. [cit. 2024-04-06].
- [4] Cobots (United States). Northwestern University. Přihl.: Oct. 28, 1997. Uděl.: Sep 14, 1999. 5,952,796. Dostupné také z: <https://patents.google.com/patent/US5952796A/en>.
- [5] PITTMAN, Kagan. A History of Collaborative Robots: From Intelligent Lift Assists to Cobots. Online. Engineering.com. 2016. Dostupné z: <https://www.engineering.com/story/a-history-of-collaborative-robots-from-intelligent-lift-assists-to-cobots>. [cit. 2024-04-06].
- [6] INSTITUTE OF ROBOTICS AND MECHATRONICS. History of the DLR LWR. Online. Dlr.de. Dostupné z: [https://www.dlr.de/rm/en/desktopdefault.aspx/tabid-12464/21732\\_read-44586/](https://www.dlr.de/rm/en/desktopdefault.aspx/tabid-12464/21732_read-44586/). [cit. 2024-04-06].
- [7] INSTITUTE OF ROBOTICS AND MECHATRONICS. LWR III. Online. Dlr.de. Dostupné z: <https://www.dlr.de/rm/en/desktopdefault.aspx/tabid-12464/#gallery/29165>. [cit. 2024-04-06].
- [8] UNIVERSAL ROBOTS. JAK SPOLEČNOST UNIVERSAL ROBOTS PRODALA PRVNÍHO KOLABORATIVNÍHO ROBOTA. Online. Universal Robots. Dostupné z: <https://www.universal-robots.com/cs/o-universal-robots/centrum-novinek/historie-kolaborativn%C3%ADch-robot%C5%AF/>. [cit. 2024-04-06].
- [9] MARKETSandMARKETS. Collaborative Robot Market. Online. Marketsandmarkets.com. Dostupné z: <https://www.marketsandmarkets.com/Market-Reports/collaborative-robot-market-194541294.html>. [cit. 2024-04-06].
- [10] Machines. Technologies. Materials. Online. 2020, roč. 14, č. 3. 2020. ISSN 1314-507X. Dostupné z:

- <https://web.archive.org/web/20200805182822/https://stumejournals.com/journals/mtm/2020/3/96.full.pdf>. [cit. 2024-04-06].
- [11] FRAUNHOFER IOSB. CoBot - Collaborative Robot. Online. Iosb.fraunhofer.de. Dostupné z: <https://www.iosb.fraunhofer.de/en/projects-and-products/cobot.html>. [cit. 2024-04-06].
- [12] ABB. ABB's collaborative robot takes the strain out of sampling at Karolinska University Laboratory. Online. ABB. 2020. Dostupné z: <https://new.abb.com/news/detail/70465/abbs-collaborative-robot-at-karolinska-university-laboratory>. [cit. 2024-04-06].
- [13] FRAVEBOT. SOLUTIONS FOR SMART GREENHOUSES. Online. 2023. Dostupné z: <https://www.fravebot.com/>. [cit. 2024-04-06].
- [14] EL ZAATARI, Shirine; MAREI, Mohamed; LI, Weidong a USMAN, Zahid. Cobot programming for collaborative industrial tasks: An overview. Online. Robotics and Autonomous Systems. 2019. ISSN 09218890. Dostupné z: <https://doi.org/10.1016/j.robot.2019.03.003>. [cit. 2024-04-06].
- [15] CROWE, Steve. Universal Robots e-Series cobots launch at Automatica. Online. The Robot Report. 2018. Dostupné z: <https://www.therobotreport.com/universal-robots-e-series-cobots-automatica/>. [cit. 2024-04-06].
- [16] UNIVERSAL ROBOTS. Naše historie. Online. Universal Robots. Dostupné z: <https://www.universal-robots.com/cs/o-universal-robots/na%C5%A1e-historie/>. [cit. 2024-04-06].
- [17] Collective data sheet. Online. In: Universal Robots. 2023. Dostupné z: [https://www.universal-robots.com/media/1829346/11\\_2023\\_collective\\_data-sheet.pdf](https://www.universal-robots.com/media/1829346/11_2023_collective_data-sheet.pdf). [cit. 2024-04-06].
- [18] The UR+ ecosystem explained. Online. In: UNIVERSAL ROBOTS. Assemblymag. Dostupné z: <https://www.assemblymag.com/ext/resources/images/2019/UR/UR-Ecosystem-Explained.pdf>. [cit. 2024-04-06].
- [19] UNIVERSAL ROBOTS. AX Series - Robotiq Palletizing Solution. Online. Universal Robots. Dostupné z: <https://www.universal-robots.com/plus/products/robotiq/ax-series-robotiq-palletizing-solution/>. [cit. 2024-04-06].
- [20] Datasheet: ur3\_tech\_spec\_web\_en. Online. In: Universal Robots. Dostupné z: [https://www.universal-robots.com/media/1828034/ur3\\_tech\\_spec\\_web\\_en.pdf](https://www.universal-robots.com/media/1828034/ur3_tech_spec_web_en.pdf). [cit. 2024-04-06].
- [21] Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods. Online. A John Wiley, 2009. ISBN 978-0-470-69683-5. Dostupné z: <https://ccc.inaoep.mx/~villasen/bib/Automatic%20Speech%20and%20Speaker%20Recognition%20Large%20Margin%20and%20Kernel%20Methods.pdf>. [cit. 2024-04-07].



- [22] DILMEGANI, Cem. Top 11 Voice Recognition Applications in 2024. Online. AIMultiple. 2024. Dostupné z: <https://research.aimultiple.com/voice-recognition-applications/v>. [cit. 2024-04-07].
- [23] JUANG, B.H a R. RABINER, Lawrence. Automatic Speech Recognition - A Brief History of the Technology Development. Online. 2004. Dostupné z: [https://www.researchgate.net/publication/249888949\\_Automatic\\_Speech\\_Recognition\\_-\\_A\\_Brief\\_History\\_of\\_the\\_Technology\\_Development](https://www.researchgate.net/publication/249888949_Automatic_Speech_Recognition_-_A_Brief_History_of_the_Technology_Development). [cit. 2024-04-07].
- [24] IBM. Speech recognition. Online. IBM. Dostupné z: <https://www.ibm.com/history/voice-recognition>. [cit. 2024-04-07].
- [25] TOTAL VOICE TECHNOLOGIES. A Brief History of Voice Recognition Technology. Online. Total Voice Technologies. Dostupné z: <https://www.totalvoicetech.com/a-brief-history-of-voice-recognition-technology/>. [cit. 2024-04-07].
- [26] IBM. Pioneering Speech Recognition. Online. The Wayback Machine. Dostupné z: <https://web.archive.org/web/20150219080748/http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/speechreco/>. [cit. 2024-04-07].
- [27] XIAO, Longwei. The Development and Status of Speech Recognition: An Overview. Online. Applied and Computational Engineering. 2023, roč. 8, č. 1. ISSN 2755-2721. Dostupné z: <https://doi.org/10.54254/2755-2721/8/20230127>. [cit. 2024-04-07].
- [28] RABINER, L.R. A tutorial on hidden Markov models and selected applications in speech recognition. Online. Proceedings of the IEEE. Roč. 77, č. 2, s. 257-286. ISSN 00189219. Dostupné z: <https://doi.org/10.1109/5.18626>. [cit. 2024-04-25].
- [29] ŽÁKOVÁ, Kristýna. Implementace rozpoznávače na bázi GMM-HMM v programovém systému MATLAB. Diplomová. České vysoké učení technické v Praze, 2020.
- [30] C. AGGARWAL, Charu. Neural Networks and Deep Learning. Online. Springer, 2018. ISBN 978-3-319-94462-3. Dostupné z: <https://doi.org/10.1007/978-3-319-94463-0>. [cit. 2024-04-27].
- [31] GEORGESCU, Alexandru-Lucian; PAPPALARDO, Alessandro; CUCU, Horia a BLOTT, Michaela. Performance vs. hardware requirements in state-of-the-art automatic speech recognition. Online. EURASIP Journal on Audio, Speech, and Music Processing. 2021, roč. 2021, č. 1. Dostupné z: <https://doi.org/10.1186/s13636-021-00217-4>. [cit. 2024-04-07].
- [32] WANG, Dong; WANG, Xiaodong a LV, Shaohe. An Overview of End-to-End Automatic Speech Recognition. Online. Symmetry. 2019, roč. 11, č. 8. ISSN 2073-8994. Dostupné z: <https://doi.org/10.3390/sym11081018>. [cit. 2024-04-07].
- [33] HALADA, Martin. Systém rozpoznávání řeči typu E2E. Online, Diplomová práce. Liberec: Technická univerzita Liberec, 2023. Dostupné z: <https://dspace.tul.cz/server/api/core/bitstreams/ae2567a0-0c38-40da-b5cd-3f9bb465d2fe/content>. [cit. 2024-04-07].

- [34] OPENAI. Introducing Whisper. Online. OpenAI. 2022. Dostupné z: <https://openai.com/research/whisper>. [cit. 2024-04-09].
- [35] OPENAI. Whisper. Online. Github. Dostupné z: <https://github.com/openai/whisper>. [cit. 2024-04-09].
- [36] IBM CLOUD. Speech to Text docs. Online. 2021. Dostupné z: <https://cloud.ibm.com/docs/speech-to-text>. [cit. 2024-04-09].
- [37] ZHANG, Anthony. SpeechRecognition 3.10.3. Online. 2017. Dostupné z: <https://pypi.org/project/SpeechRecognition/>. [cit. 2024-04-11].
- [38] AMOS, David. The Ultimate Guide To Speech Recognition With Python. Online. 2018. Dostupné z: <https://realpython.com/python-speech-recognition/>. [cit. 2024-04-11].
- [39] PYTHON. The Python Tutorial. Online. Python. Dostupné z: <https://docs.python.org/3/tutorial/index.html>. [cit. 2024-04-16].
- [40] PYPI. Ur-rtde 1.5.7. Online. Pypi. 2023. Dostupné z: <https://pypi.org/project/ur-rtde/>. [cit. 2024-04-16].
- [41] UNIVERSITY OF MANITOBA. Acoustic Phonetics: Formants. Online. 2005. Dostupné z: <https://home.cc.umanitoba.ca/~krussll/phonetics/acoustic/formants.html>. [cit. 2024-04-21].
- [42] Unimation, el primer brazo robot creado en 1961 para la industria. Online. In: Azulweb.net. 2022. Dostupné z: <https://www.azulweb.net/unimation-el-primer-brazo-robot-creado-en-1961-para-la-industria/>. [cit. 2024-04-06].
- [43] COLGATE, J.E.; PESHKIN, M. a KLOSTERMEYER, S.H. Intelligent assist devices in industrial applications: a review. Online. Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453). 2003. ISBN 0-7803-7860-1. Dostupné z: <https://doi.org/10.1109/IROS.2003.1249248>. [cit. 2024-04-06].
- [44] GIL, Jorge Juan; SÁNCHEZ, Emilio; HULIN, Thomas; PREUSCHE, Carsten a HIRZINGER, Gerd. Stability Boundary for Haptic Rendering: Influence of Damping and Delay. Online. Journal of Computing and Information Science in Engineering. 2009, roč. 9, č. 1. ISSN 1530-9827. Dostupné z: <https://doi.org/10.1115/1.3074283>. [cit. 2024-04-06].
- [45] PolyScope 5. Online. In: UNIVERSAL ROBOTS. Universal Robots. Dostupné z: <https://www.universal-robots.com/products/polyscope-5/>. [cit. 2024-04-06].
- [46] UNIVERSAL ROBOTS. Produkty. Online. Universal Robots. Dostupné z: <https://www.universal-robots.com/cs/produkty/>. [cit. 2024-04-06].
- [47] Insights from the International Conference on Acoustics, Speech, and Signal Processing. Online. In: Appen. 2019. Dostupné z: <https://www.appen.com/blog/insights-from-the-international-conference-on-acoustics-speech-and-signal-processing>. [cit. 2024-04-07].

## SEZNAM ZKRATEK A SYMBOLŮ

API – Application Programming Interface

CNN – Convolutional Neural Network

DARPA – Defense Advanced Research Projects Agency

db – decibel

DNN – Deep Neural Network

GRU – Gated Recurrent Unit

GMM – Gaussian Mixture Model

HMM – Hidden Markov Models

IoT – Internet of Things

IP – Internet Protocol

ISO – International Organization for Standardization

ISO/TS – Technical Specification

kHz – kilohertz

kg – kilogram

LSTM – Long Short-Term Memory

mm – milimetr

OSHA – Occupational Safety and Health Administration

ReLU – Rectified Linear Unit

RNN – Recurrent Neural Network

ROS – Robot Operating System

R.U.R. – Rossumovi univerzální roboti

SUR – Speech Understanding Research

USA – United States of America

°C – stupně Celsia

obr. – obrázek

tab. – tabulka

tzv. – takzvaně



## SEZNAM OBRÁZKŮ

Obr. 1: Robot Unimate [42].....	18
Obr. 2: Kobot od společnosti Cobotics [43] .....	19
Obr. 3: Robot KUKA LWR III [44] .....	20
Obr. 4: Chytré montážní pracoviště CoBot [11].....	21
Obr. 5: Kolaborativní robot YuMi od společnosti ABB manipulující se zkumavkami v univerzitní nemocnici Karolinska [12] .....	21
Obr. 6: FRAVEBOT při sběru jahod [13] .....	22
Obr. 7: Teach Pendant s uživatelským rozhraním Polyscope 5 od společnosti Universal Robots [45] .....	23
Obr. 8: Přehled všech nabízených kolaborativních robotů od společnosti Universal Robots [46].....	24
Obr. 9: Paletizační řešení od společnosti ROBOTIQ [19].....	25
Obr. 10: Kolaborativní robot UR3 CB série v robotické laboratoři na ÚAI .....	26
Obr. 11: Ilustrace využití rozpoznávání řeči [47] .....	27
Obr. 12: Závislost formantu 1 vůči formantu 2 reprezentující jednotlivá čísla (popisky přeloženy autorem) [23].....	28
Obr. 13: Pětistavový levo-pravý HMM model [29] .....	30
Obr. 14: Schéma hluboké neuronové sítě .....	31
Obr. 15: Princip funkce tradičních modulárních systémů (popisky přeloženy autorem) [31].....	32
Obr. 16: Princip funkce end-to-end systémů (popisky přeloženy autorem) [31] .....	33
Obr. 17: Vývojový diagram znázorňující princip funkce programu .....	46
Obr. 18: Uživatelské prostředí Polyscope v simulátoru URSim 3.15.8 pro sérii CB.....	48
Obr. 19: Testování v robotické laboratoři na ÚAI.....	49



## SEZNAM TABULEK

Tab. 1: Parametry všech nabízených kolaborativních robotů od společnosti Universal Robots [17] .....	25
Tab. 2: Dostupné modely Whisperu [35] .....	36
Tab. 3: Dostupné „Telephony“ modely služby IBM Watson Speech to Text [36].....	37
Tab. 4: Dostupné „Multimedia“ modely služby IBM Watson Speech to Text [36] ...	38
Tab. 5: Testovací příkazy.....	39
Tab. 6: Výsledky rozpoznání hlasových příkazů mužského respondenta pomocí modelu Whisper.....	40
Tab. 7: Výsledky rozpoznání hlasových příkazů ženské respondentky pomocí modelu Whisper.....	40
Tab. 8: Výsledky služby IBM Watson Speech To Text pro hlasové příkazy mužského respondenta.....	41
Tab. 9: Výsledky služby IBM Watson Speech To Text pro hlasové příkazy ženské respondentky.....	41
Tab. 10: Výsledky služby Google Speech Recognition pro hlasové příkazy mužského respondenta.....	42
Tab. 11: Výsledky služby Google Speech Recognition pro hlasové příkazy ženské respondentky.....	42
Tab. 12: Vyhodnocení testování.....	43





## SEZNAM PŘÍLOH

### **Elektronické přílohy:**

BP\_239056\_prilohy.zip – Archiv obsahující zdrojové kódy a videodokumentaci z testování