



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta dopravní

Ústav letecké dopravy

**Nositelné řešení pro detekci pohybu očí pro aplikace v reálném
prostředí**

**Wearable Eye Movement Detection Solution for Real-World
Applications**

Diplomová práce

Studijní program: Technika a technologie v dopravě a spojích

Studijní obor: Provoz a řízení letecké dopravy

Vedoucí práce: doc. Ing. Bc. Vladimír Socha, Ph.D.

Bc. Jakub Vídeňský

Praha 2022



K621.....Ústav letecké dopravy

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení studenta (včetně titulů):

Bc. Jakub Vídeňský

Studijní program (obor/specializace) studenta:

navazující magisterské –PL– Provoz a řízení letecké dopravy

Název tématu (česky): **Nositelné řešení pro detekci pohybu očí pro aplikace v reálném prostředí**

Název tématu (anglicky): **Wearable Eye Movement Detection Solution for Real-World Applications**

Zásady pro vypracování

Při zpracování diplomové práce se řiďte následujícími pokyny:

- Cílem práce je sestavit nositelné řešení pro detekci pohybu očí pro aplikace v reálném prostředí spolu s návrhem softwarového řešení, které by mělo zobrazovat polohu očí v zaznamenané scéně včetně zobrazení oblastí zájmů, které jsou v kokpitu letadla.
- Vypracujte analýzu současného stavu orientovanou na sledování pohybu očí pro aplikace v reálném prostředí (tedy ne s využitím virtuální nebo rozšířené reality).
- Sestrojte a ověřte funkčnost prostředku pro záznam pohybu očí a vizuálního pole.
- Vytvořte softwarové řešení, které umožní online sledování pohybu očí v zaznamenané scéně (vizuální pole).
- Vytvořte detektor objektů pro automatické rozeznání oblastí zájmu v prostředí vybraného kokpitu a toto řešení implementujte do vizualizačního softwaru spolu s možností exportu dat, která mimo jiné budou popisovat ve kterém čase subjekt sledoval danou oblast zájmu.
- Proved'te diskuzi, stanovte limitace a formulujte závěry práce.



- Rozsah grafických prací: dle pokynů vedoucího diplomové práce
- Rozsah průvodní zprávy: minimálně 55 stran textu (včetně obrázků, grafů a tabulek, které jsou součástí průvodní zprávy)
- Seznam odborné literatury: Leveson, N. Engineering a Safer World: System Thinking Applied to Safety, 2016, MIT Press Ltd.
Leveson, N., Thomas, J. STPA Handbook, 2018, MIT EUROCONTROL, Safety Assessment Methodology (SAM)

Vedoucí diplomové práce: **doc. Ing. Bc. Vladimír Socha, Ph.D.**

Datum zadání diplomové práce: **16. července 2021**
(datum prvního zadání této práce, které musí být nejpozději 10 měsíců před datem prvního předpokládaného odevzdání této práce vyplývajícího ze standardní doby studia)

Datum odevzdání diplomové práce: **16. května 2022**
a) datum prvního předpokládaného odevzdání práce vyplývající ze standardní doby studia a z doporučeného časového plánu studia
b) v případě odkladu odevzdání práce následující datum odevzdání práce vyplývající z doporučeného časového plánu studia

doc. Ing. Jakub Kraus, Ph.D.
vedoucí
Ústavu Ústav letecké dopravy



doc. Ing. Pavel Hrubeš, Ph.D.
děkan fakulty

Potvrzuji převzetí zadání diplomové práce.

Bc. Jakub Vídeňský
jméno a podpis studenta

V Praze dne.....16. července 2021



Abstrakt

Tato diplomová práce se zabývá problematikou sledování pohybu očí a klasifikací sledovaných oblastí zájmu. Cílem bylo navrhnout nositelné řešení, které bude odpovídat vysokému standartu kvality podobných zařízení a zároveň navrhnout systém pro automatickou detekci oblastí zájmu. V teoretické části jsou popsány obecné principy sledování pohybu očí, včetně přiblížení problematiky samotných pohybů očí. Za účelem pochopení problematiky automatické detekce oblastí zájmu je zde popsán princip umělých neuronových sítí a detektorů objektů. Teoretický úvod do problematiky je doplněn rešerší z prací, které se této problematice věnovaly v minulosti. Po návržení a sestrojení nositelného řešení proběhlo jeho ověření a použití v kokpitu letounu Cessna 182 a naměřená data byla zpracována pomocí detektoru objektů. V závěru jsou formulovány výsledky této práce a zároveň navrženo budoucí potenciální využití tohoto systému.

Klíčová slova: detektor oblastí zájmu, sledování pohybu očí, umělá neuronová síť



Abstract

This master's thesis studies problematics of eye tracking and areas of interest classification. The main goal of the thesis was to design wearable eye tracking device, that would meet high standard quality criteria of similar devices, and at the same time to develop solution for automatic areas of interest detection. In the theoretical part there are summarized basic principles used in eye tracking, including description of eye movements for better understanding of imposed requirements on the device. To better understand automatic area of interest detection, principles of artificial neural networks and object detectors are described in this part. The theoretical introduction to the problematics is followed with literary review of some older articles focusing on similar topic. After designing and after the production the device was tested in real conditions in the cockpit of Cessna 182. Gathered data were processed by an automatic object detector. In the last part of the thesis is the summary of the results and there is also a proposal for future potential use of this system.

Keywords: areas of interest detector, artificial neural network, eye tracking



Poděkování

Rád bych touto cestou poděkoval doc. Ing. Bc. Vladimíru Sochovi za odborné vedení, pomoc a rady při psaní této práce. Dále bych chtěl poděkovat Ing. Stanislavu Kušmírkovi za pomoc při návrhu modelu pro 3D tisk. Rád bych poděkoval také své rodině a všem přátelům, kteří mě při vytváření této práce podpořili, a bez jejichž pomoci by nebylo možné práci dokončit.



Čestné prohlášení

Prohlašuji, že jsem diplomovou práci s názvem Nositelné řešení pro detekci pohybu očí pro aplikace v reálném prostředí vypracoval samostatně a použil k tomu úplný výčet citací použitých pramenů, které uvádím v seznamu přiloženém k diplomové práci.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu §60 Zákona č.121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne 12. května 2022

Videňský
.....
Podpis



Obsah

Úvod	10
1 Analýza současného stavu	11
1.1 Pohyby očí	11
1.1.1 Sakadické pohyby	12
1.1.2 Mikrosakadické pohyby	13
1.1.3 Plynulé sledování	13
1.1.4 Vergence	15
1.1.5 Vestibulo-okulární pohyby	15
1.1.6 Optokinetický nystagmus/optokinetický pohyb	16
1.2 Metody sledování pohybu očí	16
1.2.1 Mechanické metody	17
1.2.2 Elektrookulografie	19
1.2.3 Videookulografie	22
1.3 Rešerše	25
2 Umělé neuronové sítě	33
2.1 Učení neuronových sítí	33
2.1.1 Chybové funkce	37



2.1.2	Inicializace vah	39
2.1.3	Zpětné šíření chyby	40
2.2	Aktivační funkce	41
2.3	Vícevrstvé neuronové sítě	47
2.4	Detekce objektů	50
2.5	Parciální závěr	53
3	Metody	55
3.1	Návrh experimentu	55
3.2	Zařízení	56
3.3	Příprava experimentu	59
3.4	Průběh experimentu	62
3.5	Metody zpracování dat	64
4	Prezentace výsledků	71
5	Diskuze výsledků	77
6	Závěr	80
	Seznam použité literatury	82



Seznam obrázků

1.1	Sakadický pohyb očí při čtení textu.	12
1.2	Vergence - nesynchronní pohyb očí při sledování předmětů.	14
1.3	Celooční kontaktní torzní čočka s torzní a primární cívkou.	18
1.4	Rozmístění elektrod při EOG	21
1.5	Purkyněho odrazy paprsku od jednotlivých vrstev v oku.	24
1.6	Záběr z kamer snímajících oči v infračerveném spektru.	25
2.1	Modelové znázornění umělého neuronu.	34
2.2	Vliv setrvačnosti na překonání lokálního minima	35
2.3	Závislost křížové entropie na pravděpodobnosti určení výsledku	37
2.4	Gradient sestupu funkce $F(x)$	39
2.5	Grafické znázornění aktivačních funkcí neuronových sítí.	46
2.6	Princip fungování poolingové vrstvy	49
2.7	Princip fungování detektoru objektů YOLO	53
3.1	Komerčně dostupné brýle pro sledování pohybu očí	57
3.2	Držák eye trackingových kamer	58
3.3	Brýle pro sledování pohybu očí	59
3.4	Základní choreografie pro kalibraci zařízení.	61



3.5	Diagram postupu při zpracování dat.	64
3.6	Tvorba datasetu určeného pro trénování objektového detektoru.	65
4.1	PR grafy úspěšnosti detekce oblastí zájmu.	72
4.2	PR grafy úspěšnosti detekce oblastí zájmu.	73
4.3	PR graf úspěšnosti detekce na testovacím datasetu pro objekt ADF (a). . . .	74
4.4	Vizualizace všech detekovaných oblastí zájmu.	75
4.5	Výsledná podoba detektoru objektů.	76
4.6	Případ špatných světelných podmínek, kdy detekce objektů je neuspokojivá.	76



Úvod

Tématem této diplomové práce je vytvoření uceleného systému pro detekci pohybu očí pro aplikace v reálném prostředí. Součástí tohoto systému je nositelné zařízení pro sledování pohybu očí v reálném prostředí, dále pak softwarové řešení pro vyhodnocení a zobrazení polohy očí v zaznamenané scéně, a nakonec softwarové řešení pro rozpoznání oblastí zájmu v kokpitu letadla. Systém by pak byl poměrně jednoduše aplikovatelný při budoucích pokusech, které se týkají sledování pohybu očí v reálném prostředí. Takové pokusy mají velký význam vzhledem k vlivu ergonomie v kokpitu.

Tato diplomová práce je rozdělena na teoretickou a praktickou část. Teoretická část je věnována analýze současného stavu v problematice sledování pohybu očí. V této části jsou shrnuty metody, které se pro sledování pohybů očí využívají, včetně vyhodnocení jejich výhod a nevýhod. Dále jsou zde zjednodušeně popsány pohyby očí, které jsou vykonávány za různými účely. Součástí analýzy současného stavu je pak také detailní rešerše studií a prací, které byly na toto, nebo podobné téma zpracovány. Zakončením analýzy současného stavu je parciální závěr, který shrnuje nabitě znalosti a předpoklady, týkající se problematiky práce.

Praktická část práce se zaměřuje na konstrukci zařízení, návrh prototypu a vytvoření finální podoby. Dále jsou zde detailně popsány způsoby zpracování dat, která jsou pomocí takto vyrobeného zařízení získávána. Dále se tato část zaměřuje na vysvětlení principu fungování softwaru pro automatickou detekci objektů. Poslední bod praktické části se zaměřuje na zpracování konkrétních dat získaných experimentálním měřením a jejich přípravu pro strojové učení rozpoznávání oblastí zájmu.



1 Analýza současného stavu

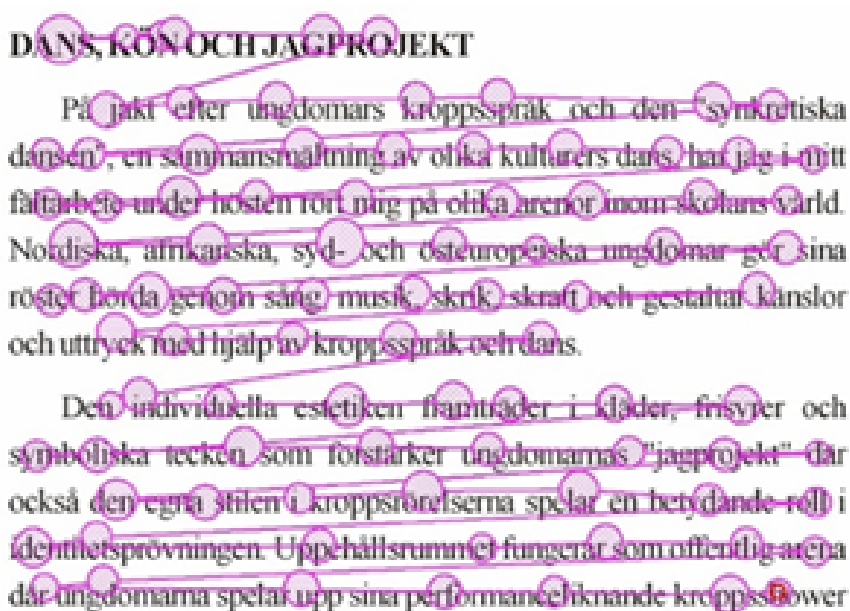
Sledování pohybu očí neboli eye tracking má dlouhou historii, ale i v dnešní době je to velmi intenzivně se rozvíjející problematika. Díky sledování pohybů očí a jejich následné analýze dokážeme získat velmi cenné informace. Pohyby očí vykonáváme jak vlastní vůlí, tak i automaticky, reflexivně. Jde o pohyby, pomocí kterých se snažíme dostat určitou oblast zájmu do zorného pole s vysokou rozlišovací schopností, a díky tomu vidět maximum detailů a získat co nejvíc informací. Sledování pohybů očí nám tak může poskytnout údaje o tom, kam se člověk dívá nebo například kudy procházela trajektorie zraku při změně oblasti zájmu. Dále můžeme analyzovat jednotlivé druhy pohybů, které vykonáváme, ať už vědomě, nebo nevědomě a detailním rozbořem těchto pohybů můžeme například diagnostikovat různá onemocnění. Díky nástupu moderních technologií se také zlepšila využitelnost sledování pohybu očí v širokém spektru aplikací. Během měření získáváme kvalitativní a kvantitativní data. Kvalitativní data zahrnují trajektorii pozorování a body zájmu, kvantitativní data zahrnují informace o délce fixace, rychlosti pohybu očí, času jednotlivých pohybů apod.

1.1 Pohyby očí

Zrakový vjem můžeme díky fyziologii oka rozdělit na dvě základní oblasti. Velmi úzká centrální část zorného pole poskytuje maximální rozlišovací schopnost, tím pádem i maximální míru detailu, který můžeme vnímat. Druhá oblast zorného pole – periferní vidění, obklopuje centrální oblast a poskytuje nižší schopnost rozeznávat detaily, která se navíc s rostoucím úhlem zorného pole snižuje. Pohyby očí jsou vykonávány z několika důvodů, hlavním důvodem je snaha dostat oblast zájmu do přímého zorného pole, s nejlepší rozlišovací schopností. Pohyby můžeme rozdělit na několik druhů v závislosti na jejich vlastnostech, jako je rychlost pohybu, vědomost či nevědomost jejich vykonávání, nebo například důvod, proč jsou vykonávány.

1.1.1 Sakadické pohyby

Sakády jsou rapidní pohyby očí, které nám umožňují velmi rychle skenovat vizuální scénu. Většinu času jsou sakády prováděny plně automaticky, bez potřeby vůle člověka, dají se však také vyvolat vůlí. Příkladem tohoto vědomého pohybu je, pokud sledujeme určitou scénu a chceme se v její blízkosti zaměřit na nějaký detail, pak se jedná právě o sakadický pohyb. Tyto pohyby se navíc vyskytují i v REM (Rapid eye movement) fázi spánku.



Obrázek 1.1: Sakadický pohyb očí při čtení textu. [1]

Čtení je obvyklý příklad sakadického pohybu očí, kdy nesledujeme plynule text, ale oči se velmi rychle pohybují po krátkých trajektoriích, jak je znázorněno na obrázku 1.1. Během čtení se navíc každé oko zaměřuje na jinou část slova, což pomáhá mozku dotvářet celkovou informaci a lépe odhadovat slova. Rozdíl ve směru pohledu obou očí se může lišit, ale většinou nepřesahuje vzdálenost dvou písmen. Tento fakt byl zjištěn právě až díky dokonalejším způsobům sledování pohybu očí. Do té doby se uvažovalo, že oči zdravého jedince fungují v perfektní synchronii, což s rostoucí přesností sledování bylo vyvráceno.



Sakadické pohyby se ale vyskytují i při fixaci na určité místo. Pohled však nezůstává na jednom bodě, ale neustále se okolo bodu zájmu pohybuje. Toto chování má původ ve fyziologii oka a sítnice. Na sítnici se nachází slepá skvrna, kde optický nerv ústí do sítnice. Pokud by se pohled nijak nepohyboval, část zorného pole by připadala právě na tuto slepou skvrnu a limitovala informace pro mozek. Tím, že se oko pohybuje i při fixaci, neustále lehce mění polohu scény, dokáže poskytnout dostatek informací a mozek je zpracuje takovým způsobem, že slepou skvrnu nevidíme [2].

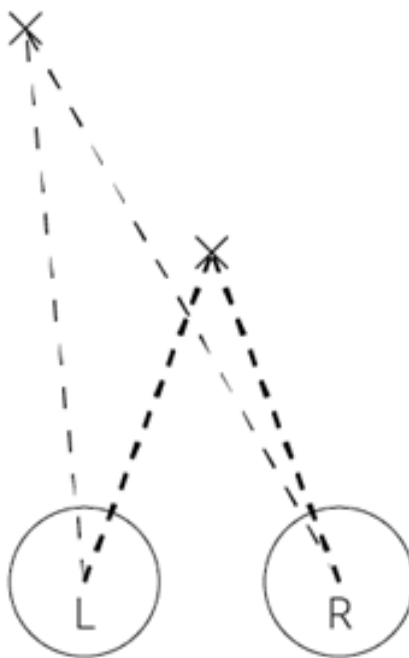
1.1.2 Mikrosakadické pohyby

Vlastností oka je, že pokud nejsou jeho zrakové receptory neustále stimulovány, postupně ztrácí citlivost a přestávají poskytovat přesné a dostatečné informace pro sestavení scény v mozku. Z tohoto důvodu oko i při fixaci na konkrétní bod neustále vykonává drobné pohyby, podobné klasickým sakadickým, ovšem s mnohem menší úhlovou výchylkou a mnohem vyšší frekvencí pohybů. Tím je zajištěno, že receptory budou neustále stimulovány novým, o kousek posunutým obrazem zorného pole. Opačným efektem stimulace je fakt, že nejme schopni vidět stíny jemných cév, které se nachází v duhovce. Tyto cévky jsou totiž vzhledem k receptorům na sítnici nehybné a nedochází tak k opakované stimulaci. Oko je schopno vykonat až 5 mikrosakadických pohybů za sekundu a jejich amplituda se nejčastěji pohybuje okolo 15 úhlových minut. Pro představu, takový úhel odpovídá přibližně šířce jednoho písmene na vzdálenosti ruky [3].

1.1.3 Plynulé sledování

Plynulé sledování vyžaduje určitý stimul, aby mohlo nastat. Minimum lidí je schopno provádět plynulý pohyb očí, aniž by docházelo ke sledování konkrétního objektu, oči se bez cíle v naprosté většině pohybují sakadicky. Pokud se ale v zorném poli pohybuje objekt, jsme schopni ho plynule pohledem následovat. Úhlová rychlost tohoto pohybu však má svá

omezení, při překročení $30^\circ/\text{s}$ již dochází k sakadickým pohybům, protože oko není schopno se plynule pohybovat s takovou úhlovou rychlostí [4]. Plynulé sledování je asymetrický pohyb, takže naše schopnost sledovat objekty se různí podle směru, kterým se pohybují. Nejvyšší rychlost můžeme dosáhnout při horizontálním pohybu objektu, naopak nejnižší úhlovou rychlost dosahujeme při pohybu objektu vertikálně ve směru nahoru. V prvotních 100 ms pohybu očí probíhá zachytávání objektu, kdy rychlost není přesně koordinována. Pokud dojde k přiblížení pohledu k objektu, srovnají se rychlosti objektu a pohledu na identické hodnoty a následuje fáze plynulého sledování v oblasti zorného pole s nejostřejším viděním. Sledování a analýza těchto pohybů se hojně využívá v lékařství, protože určité diagnózy nebo vliv léků či omamných látek postihuje právě schopnost plynulého sledování.



Obrázek 1.2: Příklad vergence – nesynchronního pohybu očí při sledování předmětů. [5]



1.1.4 Vergence

Vergence je schopnost očí, pohybovat se nesynchronně. Každé oko se může pohybovat trochu jinou úhlovou rychlostí, nebo se mohou pohybovat i mírně proti sobě nebo od sebe. Toto chování souvisí s pozorováním objektů ve 3D světě, kdy se cíle mohou pohybovat všemi směry. Na obrázku 1.2 je znázorněn princip vergence. Při pozorování blízkého objektu pohled očí směřuje ke středu souměrnosti obličeje, oči se tak musí pohybovat směrem k sobě. Při vzdálenějším objektu jsou již dráhy zraku téměř paralelní. Při pohybu například z pravé strany na levou se také každé oko musí pohybovat trochu jinou úhlovou rychlostí, aby bylo schopno sledovat cíl. Vergenci dělíme na konvergenci a divergenci. Konvergence nám umožňuje sledovat blízké objekty, zatímco divergence umožňuje sledovat objekty vzdálené [5].

1.1.5 Vestibulo-okulární pohyby

Díky vestibulo-okulárním pohybům můžeme vidět čistě a ostře i za podmínek, kdy se hlava a tělo pohybují, a tak vznikají vibrace a rušivé pohyby. Díky součinnosti vestibulárního a zrakového ústrojí se dokážou oči pohybovat takovým způsobem, že kompenzují pohyby způsobené například chůzí. Vestibulární ústrojí se nachází ve vnitřním uchu a poskytuje nám informaci o poloze těla a zrychlení ve všech směrech. Reakce na tyto vzruchy je však potřeba okamžitě, jinak by prodleva mezi pohybem těla a korekcí oka způsobovala rozmazané vidění. Takového stavu lze částečně dosáhnout konzumací alkoholu nebo jiných omamných látek, které oslabují funkci těchto pohybů. Silné vibrace mohou také způsobit narušení vidění, protože perioda je nižší než reakční doba oka, které tak reaguje na zastaralé vzruchy. Pro zajištění minimální odezvy není spojení vedeno přes mozek, ale je zde přímé spojení těchto smyslových orgánů. Reakční doba se díky tomu pohybuje okolo 10 ms, což z tohoto reflexního pohybu dělá jeden z nejrychlejších v lidském těle [6].

Pohyby, které oko při kompenzaci pohybu vykonává, dělíme do tří skupin: vertikální, horizontální a torzní. Vertikální a horizontální pohyby zastávají většinu těch, které jsou



prováděny. Torzní pohyby nedosahují takových rychlostí ani výchylek, jako zbylé druhy, ale přesto zásadně přispívají k ostrosti obrazu v náročných podmínkách.

1.1.6 Optokinetický nystagmus/optokinetický pohyb

Nystagmus nastává v okamžiku, kdy se tělo otáčí kolem své osy. Jedná se o reflexní pohyb. Vestibulární systém zaznamenává tento rotační pohyb a oči na něj patřičně reagují. Během otáčení se oči snaží kompenzovat rotační pohyb tím, že se pohybují v opačném směru. Tím umožňují krátkou fixaci na nějaký objekt. Ve chvíli, kdy objekt zmizí, vrátí se oči velmi rychle do původní pozice, a následuje další kompenzační pohyb proti směru rotace těla. Po ukončení tohoto dlouho trvajícího pohybu kolem své osy dochází k iluzi vestibulárního ústrojí a nystagmus pokračuje, i když už je rotace těla zastavena [7].

Optokinetický pohyb je podobně jako nystagmus reflexní pohyb očí. Tento pohyb nastává při sledování rychle se pohybujícího stimulu, který mizí z našeho zorného pole. Než stimul zmizí, oči ho svými pohyby následují. Jakmile zmizí ze zorného pole, oči provedou rychlý, reflexní pohyb, kterým se pohled vrátí do původní pozice, kde začalo sledování stimulu. Tento pohyb se dá sledovat například v rychle jedoucím voze, při sledování blízkého okolí. Oči se snaží zachytit nějaký konkrétní objekt a ten sledovat, než zmizí. Po jeho zmizení se velmi rychle vrací do původní pozice a sledují další stimul.

1.2 Metody sledování pohybu očí

Sledování pohybu očí má velmi dlouhou historii, první metody byly zaznamenány již v 19. století. V počátcích se jednalo pouze o pozorování oka a vyhodnocení informací, bez využití jakýchkoliv technologií. S postupem času se přistoupilo k mechanickým metodám, kdy přímo do oka byly vkládány předměty umožňující sledování a záznam dat. Takovéto metody ale nebyly uživatelsky zrovna přívětivé. Proto byly následně vyvíjeny další způsoby, jak získat přesnější data a zároveň neriskovat poškození zraku vkládáním předmětů přímo do oka



pozorovaného. Způsoby sledování pohybu očí dostupné v dnešní době tak můžeme rozdělit na 3 hlavní metody: mechanické metody, videookulografii (VOG) a elektrookulografii (EOG). Nejvíce rozvíjenou metodou je v současnosti VOG, která sama o sobě poskytuje několik možností, jak vyhodnocovat pohyby očí na základě video záznamu oka.

1.2.1 Mechanické metody

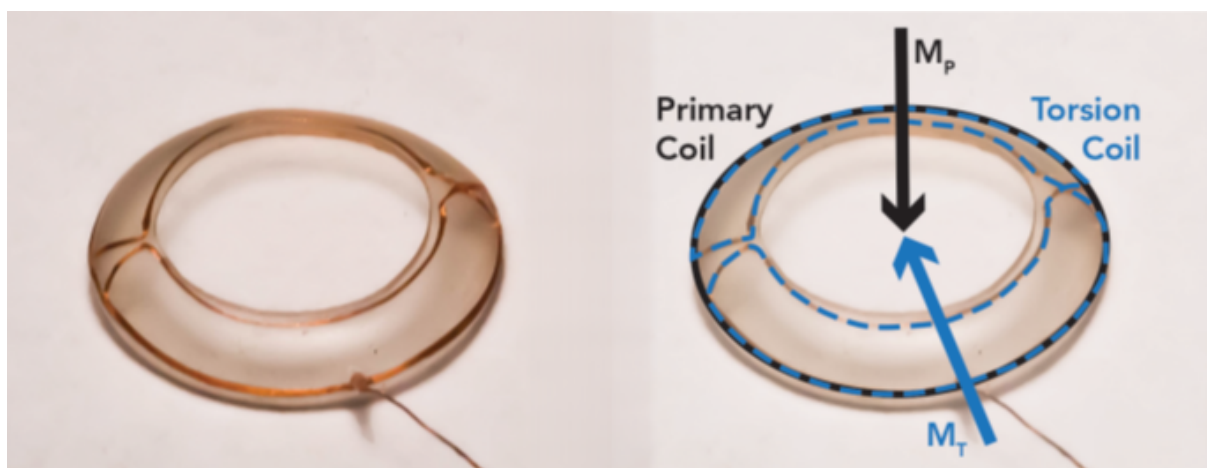
Tento přístup ke sledování pohybu očí je invazivní, součástí systému je speciální kontaktní čočka, která je vložena přímo do oka. Jedná se o celooční čočky, takzvané scleral lenses, které pokrývají celou plochu oka, včetně bělma. Právě kvůli tomu, že je takto velká čočka v přímém kontaktu se povrchem oka, působí velmi rušivě a nepříjemně. Dost často se proto měření pomocí těchto metod provádí pod částečnou anestézií oka: V minulosti se tato anestezie prováděla pomocí roztoku kokainu, vpraveného přímo na povrch oka. Mechanické metody se dají rozdělit na metody optické a magnetické.

Pro optickou metodu jsou charakteristické čočky, které jsou speciálně navrženy a vyrobeny tak, aby na jejich povrchu vznikla přesná odrazivá plocha. Během samotného měření se pak využívá paprsku infračerveného záření, které se od plošky na čočce odráží. Paprsek světla je následně zaznamenán a je vyhodnocen. Drobnou nevýhodou takového řešení je nevyhnutelný posun kontaktních čoček na povrchu oka. Během měření dochází k malému posunu, díky čemuž vznikají v měření nepřesnosti. Tyto nepřesnosti ale nejsou příliš proměnné v čase, takže s pomocí opakované kalibrace se dají eliminovat při vyhodnocování a můžeme tak dosáhnout velmi vysoké přesnosti.

Mnohem častější metodou oproti optické je ta magnetická. Během sledování jsou do očí vloženy také celooční kontaktní čočky, které obsahují jednu nebo dvě cívky z velmi tenkého drátu. Tyto čočky jsou ukázány na obrázku 1.3. Při využití jedné cívky se jedná o takzvaný směrový systém. Takové uspořádání nám poskytuje omezené informace, zaznamenáváme pouze horizontální a vertikální pohyby oka. Při využití čoček se směrovou i torzní cívkou

dokážeme navíc měřit i torzní pohyby oka. Tyto cívky jsou tenkým drátem přímo spojeny s měřícím zařízením.

Během měření se člověk musí nacházet v dokonale homogenním magnetickém poli, které je vytvářeno Helmholtzovými cívkami o průměru řádově jednotek metrů. Pokud bychom použili Helmholtzovu cívku s průměrem zhruba 1 metr, homogenní magnetické pole by mělo rozměry řádově v centimetrech, proto by bylo nemožné provádět jakákoliv smysluplná měření a je zapotřebí mnohem větších cívek. Tato měření musí probíhat se zafixovanou polohou hlavy, toho se docílí upnutím celé hlavy do specializovaného držáku, nebo například skousnutím pevně ukotveného nástavce [8].



Obrázek 1.3: Celooční kontaktní torzní čočka s torzní a primární cívkou. [8]

Během měření je pak člověk umístěn v homogenním magnetickém poli. Pokud dojde k pohybu cívky v homogenním magnetickém poli, začne se v ní indukovat napětí, které je zachyceno, ale pro další zpracování musí být ještě zesíleno pomocí zesilovače. Až poté může být zpracováno pomocí počítače. Princip funkce této metody je založen na tom, že cívky v kontaktních čočkách se chovají jako směrové antény, které při vystavení magnetickému poli indukují napětí odpovídající úhlu pootočení roviny cívky a roviny magnetického pole. Právě díky tomuto faktu jsme schopni určit přesný směr pohledu očí. Člověk se tedy nachází ve fixní pozici v homogenním magnetickém poli, které je vytvářeno velkými cívkami. Takovéto pole může být generováno pouhým jedním párem cívek, ale takovýto systém by nám nebyl schopen



poskytnou téměř žádné smysluplné informace. Místo toho se využívají dva nebo tři páry cívek, generující na sebe navzájem kolmá magnetická pole. Nejpřesnější možností je samozřejmě varianta se třemi na sebe kolmými magnetickými poli. Každé z homogenních magnetických polí je buzeno rozdílnou frekvencí nebo s posunutou fází. Díky tomu můžeme identifikovat indukované napětí na cívkách v kontaktní čočce, zjistit, které pole vyvolalo indukci daného naměřeného napětí. Při počtu dvou cívek a tří magnetických polí tedy získáváme 6 signálů, které zesilujeme, analyzujeme, a s jejichž pomocí určujeme pohyb oka.

Hlavní výhodou a důvodem, proč se mechanické metody dodnes v určitých případech používají, je jejich vysoká přesnost. Ta se pohybuje v rozmezí 5 až 10 úhlových sekund. Nejčastější aplikací se tak stává lékařství, ale také při zkoumání dynamiky pohybu oka a jeho fyziologie. Celkově by se využití dalo shrnout jako většinou vědecké a lékařské, nevyužívá se v zábavním průmyslu a komerční sféře. Nevýhodou naopak je, že provádění velmi přesných měření je velmi náročné, zejména proto, že čočky se na oční bulvě mohou nepatrně pohybovat nebo pootočít. Je proto velmi důležitá jejich správná aplikace a provedení měření. Nevýhodou využití čoček s odrazivou plochou je úzký rozptyl měření (okolo 5 stupňů) nebo jejich hmotnost, která v některých případech nemusí být zanedbatelná a může ovlivnit přesnost měření. Čočky jsou také většinou nekomfortní na nošení. Celkově se tak dají mechanické metody shrnout jako velmi přesné, zároveň je ale jejich použití velmi omezené, a proto je jejich využití možno jen ve specializovaných případech.

1.2.2 Elektrookulografie

EOG je metoda, která má oproti ostatním poměrně odlišné vlastnosti a podmínky použití. Metoda se poměrně často využívala v 70. letech minulého století, protože oproti mechanickým metodám byla levnější a dostupnější. Zároveň byla přesnější než primitivní mechanické metody. Oproti tomu videookulografie se v té době nijak zásadně nerozvíjela, zařízení sice existovala, ale byla poměrně rozměrná a těžká. Největší překážkou pak bylo vyhodnocení dat. V dnešní době se využívá vyspělých digitálních technologií, v 70. letech však vyhodnocování

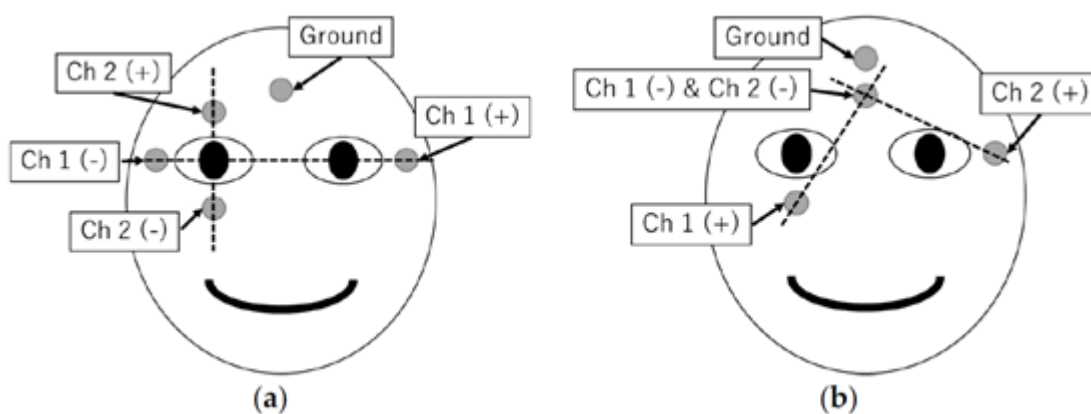


probíhalo manuálně, a proto bylo nevhodné pro většinu aplikací. I v dnešní době však EOG může najít uplatnění. Tato metoda sice nedokáže vyhodnocovat pohyby očí vzhledem k okolnímu prostředí, vyhodnocuje pohyby pouze ve vztahu k hlavě. Její hlavní výhodou však je, že pro měření není potřeba, aby měl člověk otevřené oči. Z tohoto důvodu se EOG i v dnešní době využívá k měření oční aktivity během spánku a k měření spánkových cyklů. Spánkové cykly se totiž dají dělit i podle toho, jakou oční aktivitu během nich máme. Další výhodou je pak její nízká náročnost na výpočetní výkon při záznamu a vyhodnocení a poměrně malé náklady v porovnání s magnetickou metodou.

Nevýhodou této metody je fakt, že neposkytuje tak přesné výsledky, jako magnetické mechanické metody nebo videookulografie. Není proto vhodná pro aplikace, které se zabývají interakcí s prostředím, analýzou bodů zájmu apod. Nedá se používat ani v moderních headsetech pro VR, kvůli malé přesnosti a přítomnosti elektrod. Komplikovaná je i příprava před samotným měřením, kdy je potřeba přesně umístit všechny elektrody. Signály navíc mohou být rušeny elektrickými signály, které se mohou přenášet kůží. V současné době se především schopnosti záznamu a analýzy dat posouvají směrem, který přináší velký nárůst přesnosti i pro EOG. I tato metoda tak v současnosti dosahuje velmi slušných výsledků, které se dají reálně využít.

Hlavním principem je měření změn elektrického potenciálu v kůži v blízkém okolí oka pomocí elektrod. Změna potenciálu je vyvolávána samotným pohybem oka, které se chová jako elektrický dipól s kladným pólem na rohovce a záporným pólem na sítnici. Svým pohybem poté vyvolává změny elektrického potenciálu ve svém okolí, které můžeme zaznamenat. Tato metoda je vhodná pro měření výchylky do 30 stupňů kdy je závislost velikosti amplitudy na pohybu očí lineární, nad 30 stupňů už lineární není, proto se za tuto mez nevyužívá. Frekvence signálu EOG se pohybuje od DC do 38 Hz, to představuje frekvenci mikrosakadických a sledovacích pohybů. Velikost amplitudy je přímo závislá na pohybu oka. Změny potenciálu se pohybují v rozmezí 15 až 200 μV , kde změna o 5-20 μV znamená pohyb o jeden úhlový stupeň. Velkou nevýhodou je fakt, že data jsou snadno ovlivnitelná. Mezi faktory, které dokáží signál ovlivnit patří například mrkání, pohyby obličejových svalů nebo mozková aktivita [9].

Nevýhodou EOG je pozvolný posun v měřených hodnotách, tzv. drift. Pokud probíhá měření pár desítek minut, nepředstavuje posun velký problém, pokud ale probíhá měření v řádu hodin, je nutné jednou za čas provést kalibraci změřením potenciálu při sledování přesně definovaného bodu, u kterého známe pozici vůči očím. Tím zkalibrujeme zařízení a měření může probíhat nadále přesně. Drift vzniká z mnoha důvodů. Jedním z nich je ovlivnění vodivosti kontaktu pokožky s elektrodou potem, EEG gely jsou však proti tomuto rušení poměrně rezistivní. Drift může také vznikat vlivem rušivého pozadí měření nebo polarizací elektrod. Drift neovlivňuje měření u sakadických pohybů, má však zásadní vliv při měření všech ostatních pohybů očí. Rušení vzniklé mrkáním se v signálu projevuje jako ostrá potencionálová špička, kterou ale můžeme odfiltrovat.



Obrázek 1.4: Rozmístění elektrookulografických elektrod (a) v uspořádání plus „+“ (b) v uspořádání křížem „X“. [9]

Elektrody zajišťují vodivé spojení mezi pokožkou a měřícím přístrojem. K přenosu s co nejnížší ztrátou a rušením se využívá gelu, který zprostředkovává vodivé spojení kůže a elektrody. V základu dělíme umístění elektrod pouze na dvě varianty. Varianta „X“ nebo „+“ zapojení. Obě varianty se skládají z dvou párů elektrod a jedné referenční elektrody. V případě zapojení „+“ se nachází elektrody pro snímání horizontálních pohybů vedle očí na vnějším okraji obličeje a elektrody pro snímání vertikálního pohybu jsou nad a pod jedním okem. Referenční elektroda se je umístěna zhruba uprostřed čela a poskytuje pro měření referenční hodnotu elektrického potenciálu z místa, které už není pohybem očí ovlivněno. K



tomuto bodu se pak vztahují rozdíly naměřené na jednotlivých elektrodách. Přesnost tohoto umístění elektrod se pohybuje v horizontální rovině okolo 1 stupně a vertikálně zhruba 1,5 stupně. V případě zapojení křížem se nachází jedna společná elektroda pro oba kanály nad úrovní očí v ose obličejce. Opačná elektroda jednoho kanálu se pak nachází vedle oka na vnějším okraji obličejce a opačná elektroda druhého kanálu se nachází pod druhým okem. Referenční elektroda se nachází opět na čele, nad společnou elektrodou obou kanálů. Přesnost tohoto rozvržení elektrod je v horizontální rovině zhruba 2,5 stupně a ve vertikální rovině zhruba 2 stupně [9].

1.2.3 Videookulografie

V dnešní době je videookulografie nejrozšířenější metodou pro sledování pohybu očí. Její popularita mezi ostatními metodami spočívá v několika zásadních výhodách. V první řadě je to metoda v porovnání s ostatními poměrně levná, protože nevyžaduje mnoho speciálního vybavení, ale také je neinvazivní, takže se dá používat v mnohem větším spektru aplikací. K jejímu využití je potřeba pouze videokamera a počítač, který slouží k záznamu dat a jejich následnému vyhodnocení. Metod, které se obecně nazývají videookulografie je několik, všechny je spojuje fakt, že jsou založeny na snímání oka pomocí videokamery. Přesto se od sebe tyto metody liší v několika aspektech, jako například ve spektru elektromagnetického záření, které snímají, nebo v konkrétní oblasti oka, kterou snímají za účelem určení směru pohledu. Sledované oblasti pak nejčastěji můžeme rozdělit na 3 základní, sledování zornice, sledování rozhraní mezi duhovkou a bělmem, nebo pozorování odrazů paprsků dopadajících na jednotlivé vrstvy oka (Purkyněho obrazce).

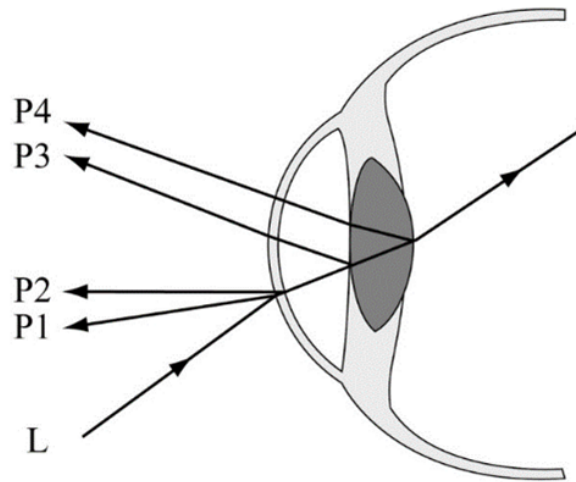
Samotné snímání oka nám poskytuje informace o pohybech očí, ale neposkytne nám informace o směru pohledu. Pro určení pohledu je proto potřeba nějaké další zařízení, které zajistí propojení dat o pohybech očí a jejich polohu vzhledem k reálnému prostředí. Jednou z metod je využití čelní kamery, která snímá okolní prostředí ve směru zorného pole člověka. Díky procesu kalibrace dojde k sladění informací o pohybech očí ve vztahu k prostředí a systém na



základě kalibrace dokáže určovat výsledný směr pohledu. Přesnost určení pohledu pak záleží na kvalitě kalibrace, celkovému posunu zařízení na hlavě v průběhu měření, nebo samotné konfiguraci zařízení kdy záleží, jestli je vybaveno jednou, nebo více kamerami. Další možností pro určení výsledného pohledu je externí snímání polohy hlavy, kdy kamera umístěná před člověkem celou dobu snímá jeho obličej a opět na základě kalibračních dat dokáže dopočítat směr pohledu v reálné scéně. Nevýhodou tohoto řešení je to, že musí být stacionární, případně s malou výchytkou v rámci zorného pole kamery. Systém s čelní kamerou snímající vizuální pole disponuje větší schopností mobility.

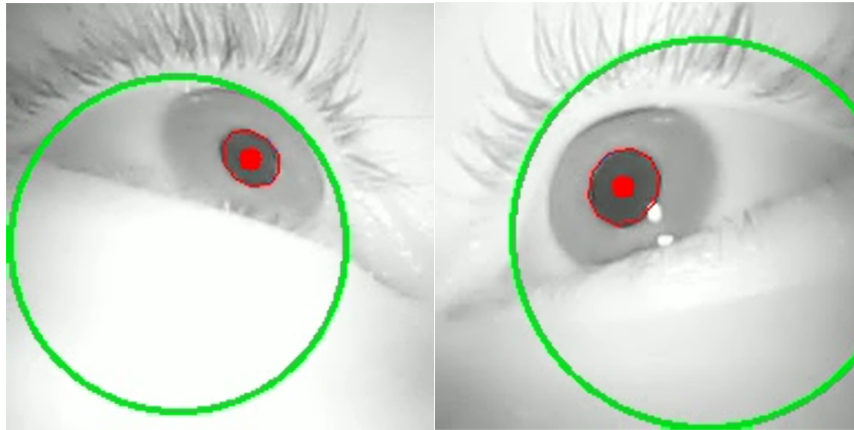
Podle oblasti oka, kterou snímáme můžeme sledování pohybu očí dělit na 2D a 3D. Pokud sledujeme zornice, jedná se o 2D metodu. Zornice je velmi jasně ohraničený útvar, který je snadno detekovatelný, ale neposkytuje nám dodatečné údaje o torzních pohybech oka. Tato metoda poskytuje pouze údaje o horizontální a vertikální poloze zornice oka. Pro získání informací o torzních pohybech oka je proto potřeba sledovat ještě hranici bělma a duhovky, protože jemné žilkování duhovky nám poskytuje dodatečnou informaci o torzi oka. Většina zařízení, která snímají oči, pracují v infračerveném spektru záření. To s sebou přináší především výhodu v podmínkách s nedostatkem světla. Infračervené kamery dokážou fungovat i za úplné tmy, pro svoji funkci využívají přisvětlovací infračervené diody. Zařízení mohou být také navržena jako monokulární, nebo binokulární. Obvyklá snímkovácí frekvence se pohybuje mezi 60-120 Hz, což postačuje pro detekci většiny pohybů. Pokud je však cílem pozorování těch nejjemnějších a nejrychlejších pohybů očí, je potřeba zařízení s obnovovací frekvencí aspoň 200 Hz. [10]

Zařízení jsou velmi citlivá na pohyb kamery vůči očím během měření. Některé moderní systémy dokážou vytvářet ke každému snímku zároveň i 3D model oka, Díky tomu jsou více odolné proti nepřesnostem vznikajícím samotným používáním zařízení. Drobným pohybům na hlavě se totiž nedá vyvarovat nikdy, pokud ale dojde k většímu posunutí, přesnost dat je značně zhoršená. Eye trackery, které nejsou umístěny přímo na hlavě, vyžadují pevné uchycení hlavy během měření, aby se zabránilo posunu očí vůči kamerám. Řešením je přidání zařízení pro sledování pohybů hlavy nezávisle a následně výpočtem určit výsledný pohled.



Obrázek 1.5: Purkyněho odrazy paprsku od jednotlivých vrstev v oku. [11]

Další možností, jak sledovat pohyby očí, je pozorování Purkyněho obrazců. Tyto obrazce vznikají na povrchu jednotlivých vrstev oka, viz obrázek 1.5. Jedná se o odrazy infračerveného paprsku, který je směřován do oka diodou. Celkem můžeme zaznamenávat čtyři odrazy, postupně od vnější strany rohovky, následovaný odrazem od vnitřní strany rohovky. Tyto odrazy se označují P1 a P2. Třetí odraz P3 vzniká na vnější straně čočky a poslední odraz P4 vzniká na vnitřní straně čočky. P1 a P2 jsou si tak blízké, že se navzájem v podstatě překrývají, proto se souhrnně označují pouze jako P1. Tento odraz je nejsilnější, a proto slouží jako referenční odraz, vůči němuž se porovnává poloha odrazů dalších. Odraz P3 je virtuální, zvětšený a značně rozptýlený. Odraz P4 je pak nejslabší, ale přesto se právě tento využívá pro porovnání s P1 a určení směru pohledu. Princip spočívá v tom, že během kalibrace se určí úhly mezi jednotlivými odrazy při pohledu na předem známá místa a následně je možné dopočítat směr pohledu na základě znalosti úhlu mezi odrazy [10]. Nejčastější metodou je ale v dnešní době sledování polohy zornic. Na obrázku 1.6 je zobrazen snímek z kamer snímajících oči v infračerveném spektru.



Obrázek 1.6: Záběr z kamer snímajících oči v infračerveném spektru.

1.3 Rešerše

V následující části práce se zaměřím na nejrůznější studie a práce, které již byly v této oblasti vypracovány. Tyto práce se soustředí na problematiku sledování pohybu očí, liší se však v cílech, kterých chtěli jejich autoři dosáhnout. Dále se liší i prostředím, ve kterém byly vykonávány. Všechny práce se však týkají pouze problematiky letectví a jsou prováděny buď na simulátorech, nebo přímo za letu. Práce byly vytvořeny v různých dobách, včetně druhé poloviny 20. století, což poskytuje i náhled do praktické části měření. Můžeme tak poznat, jaké metody byly využívány před rozšířením moderních videookulografických metod a také jak byla data vyhodnocována bez možnosti velkého výpočetního výkonu, kterým disponujeme dnes.

What Do Pilots See in Displays?

V roce 1980 byla publikována práce výzkumným centrem NASA v Langley. Tato práce se zaměřovala na rozložení přístrojů a na informace, které jsou pilotům zobrazovány během letu. Pokusy probíhaly na simulátoru jednomotorového hornoplošníku spadajícího do kategorie General Aviation. Jednalo se o obecný simulátor, který byl založen hlavně na principech



obecně využívaných v kokpitu, ale nereprezentoval žádný konkrétní model. Autoři si v úvodu práce pokládají otázku, jaké informace pilot pro úspěšně pilotování letadla potřebuje. Zatímco trajektorie pohledu pilotů byla již v té době dobře známá, nebylo jasné, jakou informaci z přístrojů odečítají. Zda je zajímavá pouze obecný směr ručičky ukazatele, například že ukazuje vpravo, nebo zda je zajímavá, že se nachází například mezi hodnotami 50 a 60, nebo jestli se zaměřují na přesnou hodnotu, kterou jim daný přístroj sděluje. Pouhé dotazníkové šetření se tak rozhodli podpořit exaktním měřením, prováděným na simulátoru s pomocí okulometru.

Pro potřeby provedení experimentu bylo využíváno digitálních počítačů, které se staraly jak o zobrazování simulovaných hodnot pomocí přístrojů, tak jejich zaznamenávání pro následné vyhodnocování. Také záznam pohybu očí byl prováděn pomocí digitálního počítače. Veškeré informace byly zaznamenávány 32x za sekundu na magnetický pásek. Zařízení pro zaznamenávání pohybu očí se nazývalo okulometr a fungovalo na principu velmi základní videookulografie. Do oka byl pomocí speciální infračervené lampy směřován paprsek. Sítnice oka má velmi reflexivní vlastnosti pro infračervené světlo, které se tak od ní efektivně odráží. Tento odražený paprsek putuje zpět do elektro-optického zařízení, kterým byla na infračervené světlo citlivá televizní kamera. Tento odraz se na záznamu jevil jako bílý kruh na černém pozadí s jasným ohraničením. Dalším odrazem, který na záznamu byl patrný je odraz paprsku od rohovky. Tento odraz se jevil jako velmi světlá skvrna. Na základě porovnání vzájemné pozice středu zornice a odrazu rohovky pak bylo možné dopočítat úhlovou výchylku oka a tím pádem i směr pohledu. Takto počítaná data měla analogový charakter v podobě přesně definovaných výkyvů napětí, ale s pomocí počítače byla digitalizována a nahrávána na magnetický pásek společně s daty ze simulátoru. V rámci práce pak bylo analyzováno rozložení a pozornost na následujících 9 základních přístrojů: výškoměr, rychloměr, digitální výškoměr, ukazatel otáček motoru, zatačkoměr, směrový indikátor, ukazatel vertikální rychlosti, indikátor odchylky kurzu – CDI a indikátor sestupové roviny a localizeru.

Autoři práce analyzovali, jaká procentuální zastoupení mají jednotlivé oblasti zájmu a jaký význam má délka fixací na tyto oblasti. Původním cílem bylo navržení nového rozložení



přístrojů v kokpitu tak, aby pilot mohl jednodušeji získávat informace o letu. Již ze dřívějších studiích totiž vyplynulo, že piloti se soustředí především na výškoměr. Z tohoto přístroje nejčastěji odhlédnou na jeden další přístroj, a pak svůj pohled opět vrací na výškoměr. Záměrem této práce proto bylo vytvořit jeden přístroj, který by obsahoval několik informací a pilot tak nemusel přesouvat svou pozornost mezi jednotlivými přístroji, případně zavést určitou formu digitalizace přístrojů. Z následných měření, a především komentářů pilotů ale vyšlo najevo, že jimi takto navržený přístroj naopak snižuje výkon pilotů. Ti si stěžovali, že přístroj musí více „číst“, zatímco v klasickém rozložení ve tvaru T jim stačí se na přístroj velmi rychle podívat a mají ihned představu o informaci, kterou jim ukazuje. Jak již dnes víme, toto rozložení se v obchodní letecké dopravě téměř nevyskytuje, veškeré základní informace jsou sdruženy ve velmi přehledných displejích, nicméně v oblasti obecného letectví se stále velmi často vyskytují analogové přístroje, u kterých neproběhla zásadní snaha o sjednocení do jednoho hlavního přístroje. [12]

Efekty komplexity mise na mentální zátěž pilota, jeho výkon a situační povědomí

Tato práce byla publikována v roce 2009 zástupcem švédské výzkumné agentury pro obranu, Erlandem Svenssonem, a zástupcem výzkumné laboratoře U.S. Air Force na základně Wright-Patterson, OH, Glennem Wilsonem. Práce se zaměřovala na vliv složitosti mise, kterou piloti provádí, na jeho mentální zátěž, situační povědomí a lidskou výkonnost. Cílem bylo popsání závislosti pomocí modelu. Práce je ve své experimentální fázi rozdělena na dvě části. V první části jsou prováděna měření v reálných letounech Saab 37 Viggen. Celkem se této fáze zúčastnilo 20 aktivních pilotů švédského královského letectva, kteří takto vykonali 150 cvičných misí. Z tohoto počtu jich bylo celkem 144 vyhodnoceno a zahrnuto do hodnocení první fáze. V této fázi probíhalo poze dotazníkové šetření pilotů před a po každém letu, kdy piloti hodnotili náročnost misí pomocí několika vybraných standardizovaných dotazníků.

V další fázi probíhalo měření na simulátoru. To umožnilo implementaci mnoha měřících zařízení na sledování fyziologických projevů pilotů. Této části se účastnilo 15 pilotů a provedli



celkem 40 simulovaných misí. Celkem bylo z tohoto počtu vyhodnoceno 35 misí. Součástí získávání dat a informací byly ty stejné dotazníky, které vyplňovali i piloti reálných letových misí. Piloti je taktéž vyplňovali vždy před a po provedení mise. Zároveň bylo během této fáze sledováno několik dalších projevů, jako například frekvence srdečního tepu, frekvence mrkání, nebo délka fixací na dané přístroje a oblasti kokpitu. Míra fixací byla zvolena jako hlavní ukazatel vizuálního chování pilotů. Míra fixací v tomto případě vyjadřuje počet změn fixací za určitou časovou jednotku, zde 30 sekund.

Cílem práce bylo pochopení závislosti sledovaných parametrů na různých situacích, do kterých se piloti během pokusů dostávali. Výsledky indikují, že není znatelný rozdíl v návaznosti na rozdílné prostředí, kde se pokus konal, ať už se jednalo o simulátor, nebo reálnou situaci. Zároveň však výsledky práce ukázaly, jak jsou na sobě jednotlivé aspekty závislé. S rostoucí komplexitou informací, které jsou pilotovi zobrazovány na displeji, zároveň klesá jeho mentální rezerva pro vnímání dalších podnětů. S tím, jak roste mentální zátěž pilota, se nejprve nesnižuje situační povědomí, ani jeho výkonnost, ale s rostoucí mentální zátěží klesá schopnost udržovat si tyto aspekty na vysoké úrovni. Je proto velmi podstatné, jak jsou pilotům na displejích předávány informace. V případě bojových podmínek, nebo jakýchkoliv podmínek vyžadujících vysokou mentální kapacitu, je proto velmi vhodné snížit množství informací předávaných pilotovi na úplné minimum. [13]

Registrace a analýza pilotovy pozornosti pomocí mobilního eye trackeru

Tento výzkum navazuje na předchozí práci, ve které autoři využívali statické zařízení pro sledování pohybu očí. Takové řešení však přináší mnohé limitace, které jsou z velké části eliminovány při využití mobilního zařízení přímo na hlavě pilota. Měření s pomocí tohoto zařízení probíhala v simulátoru letounu Piper Seneca II na letišti v polském Řešově. Pro měření pohybu očí pak bylo využito zařízení Tobii Glasses Pro, které pracuje s kamerami v infračerveném spektru světla a dosahuje vynikající přesnosti. Jedná se o jeden z nejlepších systémů na trhu. Cílem práce je hodnocení možností měření a následného zpracování dat,



zároveň jsou ale naměřená data podrobena analýze, jak je mezi ně distribuována pozornost pilota a jak dlouhé jsou fixace pozornosti na tyto přístroje.

Pokusu se účastnilo celkem 10 pilotů, z nichž polovina měla IFR licenci a druhá polovina byla kvalifikována pro VFR létání. Piloti se lišili nejen danou kvalifikací, ale také zkušenostmi s daným typem létání, počtem nalétaných hodin, ale také věkem. Úkolem těchto pilotů pak bylo provedení přesného přiblížení podle postupů ILS s horizontálním i vertikálním vedením. Začátek experimentu byl zvolen ve vzdálenosti 7 NM od letiště ve výšce 3000 ft. Cílem bylo dosažení rozhodné výšky pro přistání 200 ft, samotné přistání nebylo předmětem měření. Podmínky viditelnosti byly 550 m dráhové dohlednosti a základna oblačnosti se pohybovala 200 ft nad povrchem země. Pilotům byly po dobu letu poskytovány informace z 8 přístrojů, kterými byly: umělý horizont, rychloměr, výškoměr, RMI indikátor, VOR/ILS indikátor, variometr, koordinátor zatačky a CDI indikátor. Vyhodnocení měření se pak zaměřovalo na všechny přístroje, které měli piloti k dispozici.

Výsledky práce se dají rozdělit do dvou částí na ty týkající se zařízení a detekce oblastí zájmu pilotů a na samotné vyhodnocení naměřených dat. Zařízení Tobii Glasses Pro fungovalo podle předpokladů, poskytovalo přesné informace a výsledkem tak bylo jednoznačné doporučení pro využití při podobných aplikacích. Společnost Tobii ke svému zařízení dodává i software pro zpracování dat, který obsahuje i možnost označení oblastí zájmu a následně hodnocení kvalitativních a kvantitativních ukazatelů týkajících se pohledu na tyto oblasti. Autoři však narazili na problém, že označení oblastí zájmu je statické, tudíž při označení na prvním snímku záznamu vše odpovídá zadaným parametrům, ale následně při sebemenším pohybu hlavy označené oblasti zájmu přestávají korespondovat s reálnou předlohou v podobě přístrojů. Tento problém nemohl být vyřešen manuální analýzou dat, protože to je příliš časově náročné a neefektivní, autoři proto využili separátní analýzy pomocí jiného softwaru, kdy sledovali jednotlivé oblasti zájmu zvlášť a přiřazovali fixace k daným oblastem zájmu.

Výsledky vyhodnocení naměřených dat pak přinesly některé nečekané výsledky. V případě VFR pilotů bylo zjištěno, že nejvíce pozornosti věnují v následujícím pořadí: umělému



horizontu, VOR/ILS indikátoru, rychloměru, variometru a výškoměru. V případě IFR pilotů byla tato skupina přístrojů následující: umělý horizont a VOR/ILS indikátor (podobný počet fixací), variometr, výškoměr a rychloměr. Při vyhodnocování výsledků však bylo zjištěno, že dochází k poměrně vysokému množství fixací mimo definované oblasti zájmu, tudíž pozorované výsledky mohou být ovlivněny chybou ve zpracování dat. Může také dojít k nenaplnění parametrů pro klasifikaci fixací v dané oblasti zájmu, což snižuje zaznamenaný počet fixací oproti realitě. Samotné zařízení poskytuje přesná data a chyby vznikají až při vyhodnocení dat. Autoři však na veškeré limitace upozorňují a zmiňují, že i přes tyto nedostatky má systém velký potenciál využití. [14]

Strategie a výkon pilotů při monitorování v automatické fázi letu

Tato práce vznikla součinností univerzity Ann Arbor v Michiganu, MI, univerzity Urbana-Champaign v Illinois, IL a společnosti Boeing v Seattlu, WA v roce 2007. Cílem práce bylo detailní prostudování chování pilotů při monitorování stavu letounu, parametrů a přístrojů na vysoce automatizovaných palubách moderních letadel. V průběhu několika let po přelomu tisíciletí se objevovaly indicie, že je značné riziko, že piloti během automatizované fáze letu nedokážou dodržet automatizované strategie monitorování parametrů a tím pádem nastane riziko snížení jejich situačního povědomí a výkonnosti.

Experimentu se účastnilo celkem 20 mužských pilotů. Jednalo se o piloty letadla Boeing 747-400, z nichž 10 mělo hodnost kapitána a 10 hodnost prvního důstojníka. Jednalo se o piloty dvou amerických aerolinek. Jejich nálet se pohyboval v rozmezí 100 až 9000 letových hodin na daném typu letadla. Zároveň však měli všichni alespoň 1000 hodin náletu v kokpitu s obrazovkami pro zobrazování letových, motorových a navigačních údajů, tzv. glass cockpit. Každý z účastníků před měřením poskytl své demografické údaje, které mohou posloužit pro lepší zasazení jejich výkonů do kontextu jejich osobnosti. Měření se pak účastnil každý pilot na pozici, která mu normálně přísluší. Pozici druhého pilota v kabině zastával nezávislý pilot, který byl ovšem pouze ve funkci pilota monitorujícího. Zároveň se tento pilot nijak aktivně



nepodílel na zvládnutí situací zahrnutých v daném scénáři. Jeho úkolem nebylo ani vytváření dalších problémů v kokpitu, jednalo se pouze o osobu monitorující let a doplňující posádku, jako za normálních operací. Pro potřeby experimentu byl přítomen i člověk zastupující funkci řídicího letového provozu, se kterým piloti v průběhu scénáře komunikovali.

Experiment probíhal na stacionárním simulátoru letadla Boeing 747-400 s vnějším zobrazovacím zřízením ESIG 3350 od společnosti Evans and Sutherland. Jako zařízení pro sledování pohybu očí sloužil ALS (Applied science laboratory) Series 4000 přenosný eye tracker. Toto zařízení poskytuje pouze údaje o pohybech očí na základě polohy zornice, nedokáže navíc sledovat pohyby hlavy a tím určit výsledný směr pohledu zasazený do reálné scény ve směru pohledu. V průběhu experimentu byly pilotům kladeny otázky týkající se automatických systémů letu. Cílem bylo zjistit, jak hluboce piloti chápou fungování těchto systémů a jestli například rozumí indikacím, které jim tyto systémy poskytují.

Celkem byli všichni piloti podrobeni 12 scénářům, které zahrnovali všechny části letu od vzletu, manévrování, ustáleného letu i přistání. Měření všech scénářů pak zabralo přibližně 1 hodinu. Během samotných scénářů měli piloti při automatizované fázi letu provádět veškeré standartní procedury. Během scénářů pak byl pokus několikrát zastaven a piloti byli dotazováni na aktuální stav systému automatizace letu, na aktuální módy systémů a podobné informace. Tyto dotazy ve výsledku pomáhají utvořit ucelený obraz o tom, jak efektivně piloti monitorují stav letadla. V důsledku chybné interpretace stavu systémů pak může dojít k nesprávným vstupům ze strany pilota, nebo nepochopení určitého chování letadla vlivem nevhodného zacházení se systémy automatizace letu.

Cílem práce bylo zjistit, jestli piloti věnují pozornost správným směrem ve chvíli, kdy je to potřeba. Například při indikaci změny módu navigace letadla se zkoumalo, jestli pilot vůbec zaregistruje změnu na displeji. Dílčím cílem této práce bylo zjistit, jak je distribuovaná pozornost pilotů v kokpitu během letu a také jak pilot pracuje se získanými informacemi, jestli je chápe dostatečně do hloubky a dokáže si vytvářet souvislosti mezi údaji, které vidí před sebou. Výsledky měření ukázaly, že 31 % času piloti průměrně ve všech fázích letu sledují



PFD (Primary Flight Display) na kterém se zobrazují základní informace o letu, rychlost, výška, vertikální rychlost, umělý horizont a je zde také indikace automatických systémů letadla. 25 % času sledují ND (Navigační displej) s mapou letu a informacemi o trase, a během letu pouhá 3 % času sledují okolí letadla. Toto procento se s pozdější fází přiblížení zvyšuje přibližně na 12 %. Ve zbylém čase monitorují panel autopilota nebo jiné systémy a panely v kokpitu. [15]



2 Umělé neuronové sítě

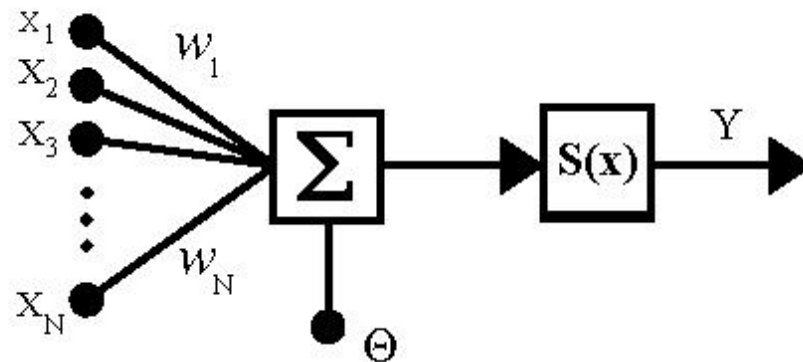
Umělé neuronové sítě jsou architektury výpočetních struktur, které se využívají v oblasti umělé inteligence. Inspiraci pro architekturu těchto sítí představuje mozek a jeho neurony. Konkrétně jde o to, jak jsou na sebe jednotlivé neurony vázány, a jakým způsobem reagují na signály ostatních neuronů. Takto navržené umělé sítě dokážou plnit úkoly, na které nebyly specificky naprogramovány, díky využití technik učení vzorců z velkých množství dat.

Základní jednotkou pro vytvoření umělé neuronové sítě je neuron. Princip fungování takového umělého neuronu je znázorněn na obrázku 2.1. Základem je vektor vstupů a vektor vah, definující důležitost každého vstupu. Odpovídající hodnoty vstupu a váhy daného vstupu jsou mezi sebou vynásobeny a všechny takto násobené hodnoty jsou následně sečteny. K sumě součinu vstupního vektoru a vektoru vah může být ještě přičtena skalární hodnota bias. Ještě před vstupem do aktivační funkce je definována prahová hodnota neuronu, která zajišťuje aktivaci, nebo neaktivitu neuronu. Pokud je totiž součin vstupního vektoru a vektoru vah nižší než daná prahová hodnota, je neuron neaktivní a nevyšle dál žádný signál. Pokud je však součin vyšší než prahová hodnota, vstupuje do aktivační funkce. Výpočet z aktivační funkce je pak jediným výstupem daného neuronu, který je buď vyslán do další vrstvy neuronů, nebo pokud se jedná o finální vrstvu neuronů, je zpracován na výsledek.

2.1 Učení neuronových sítí

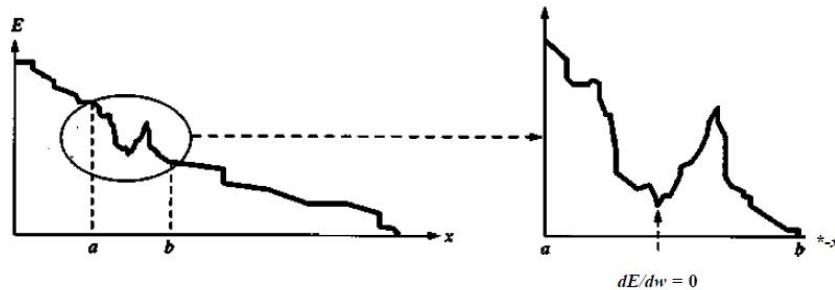
Učení umělých neuronových sítí spočívá v adaptaci sítě tak, aby lépe a přesněji dokázala zpracovávat data a poskytovala výsledky na základě prováděných pozorování vstupních dat. Učení je založeno na upravování váhových vektorů umělých neuronů, nebo změně prahové hodnoty neuronu s cílem dosažení lepších výsledků. Proces učení se dá považovat za ukončený ve chvíli, kdy další iterace již nepřináší téměř žádné, nebo vůbec žádné zlepšení přesnosti výsledků. I po dokončení učení však drtivá většina neuronových sítí nedosahuje nulové

chybovosti. Pokud je však výskyt chyb příliš vysoký, je potřeba změnit nějakou část algoritmu v závislosti na typu učení a neuronové síti, která je využívána.



Obrázek 2.1: Modelové znázornění umělého neuronu. [16]

Dalším důležitým aspektem, který hraje velkou roli při učení systému je rychlost učení. Přesněji můžeme rychlost učení definovat jako velikost korekčních kroků při jednotlivých iteracích. Čím vyšší je hodnota, o kterou se pokaždé mění váhové vektory nebo prahové hodnoty neuronů, tím rychleji se celý systém dokáže vytrénovat na daný úkol. Toto je však spojeno s nižší úrovní přesnosti výsledků. Naopak čím drobnější jsou kroky, o které se mění hodnoty při jednotlivých iteracích, tím pomaleji se systém učí, ale dosahuje vyšší přesnosti výsledků. Ve snaze najít ideální kompromis rychlosti a přesnosti se využívají neuronové sítě s proměnnou rychlostí učení. Takovýto systém dokáže měnit velikost změny důležitých hodnot a je založen na principu setrvačnosti. Velikost změny při jedné iteraci je tak závislá na velikosti změny při předchozí iteraci. Na obrázku 2.2 je znázorněn problém, který by potenciálně mohl vzniknout, kdyby změny rychlosti učení nebyly vázány principem setrvačnosti. Snahou algoritmu je hledání minima jeho chybovosti, ale algoritmus umí hledat hlavně lokální minima. Pokud jsou navíc tato minima ohraničena velkou výchytkou v chybovosti, jako je znázorněno na obrázku, tak algoritmus toto lokální minimum považuje za globální minimum. Díky setrvačnosti má ale možnost „překonat“ následující vyšší chybovost a dále pokračovat v úpravách váhových vektorů a prahových hodnot neuronů v síti s cílem najít globální minimum chybovosti.



Obrázek 2.2: Díky setrvačnosti změn vektoru vah dokáže algoritmus překonat lokální minimum. [16]

Zjednodušeně můžeme učení umělých neuronových sítí rozdělit na tři základní přístupy: učení s učitelem, učení bez učitele, a tzv. zpětnovazební učení. Každý z těchto přístupů je vhodný pro jiné aplikace. Učení s učitelem spočívá v přiřazování vstupů a výstupů na základě předem známých párů vstup a výstup. Algoritmus se učí na základě širokého datasetu označených dat, která přesně popisují žádané výsledky na základě vstupů. Může se jednat například o označené oblasti zájmu na obrázku reálné scény. Nejčastějším využitím metody učení s učitelem je klasifikace objektů a oblastí zájmu, případně určení regresní funkce.

Další metodou strojového učení umělých neuronových sítí je učení bez učitele. Tento algoritmus je schopen učení se bez předem označených dat a spoléhá na vlastní rozpoznávání vzorců ve vstupních datech. Tento typ učení se dá přirovnat k učení novorozenců. Novorozenci nejsou schopni dostatečně vnímat zpětnou vazbu, přesto v prvních měsících života začínají reagovat na určité vzorce a vlastnosti vizuálních vjemů. Tyto informace o okolním prostředí však nezískávají učení od ostatních lidí v jejich okolí, ale vlastním poznáváním světa okolo sebe. Na základě těchto biologických procesů funguje i metoda učení bez učitele, kdy se umělé neuronové sítě snaží heuristicky vyhledávat ve vstupních datech speciální vlastnosti a vzorce bez vnější pomoci. Nejčastějším využitím této metody učení je shlukování dat



(clustering), odhadování statistických funkcí rozložení dat, filtrování dat, nebo kompresní funkce při zpracování dat.

Posledním ze základních typů učení neuronových sítí je učení s pomocí zpětné vazby. Zpětná vazba je získávána od samotných pokroků během učení. Cílem této metody je dosažení výsledku s co nejnižší kumulovanou hodnotou nákladové funkce, jednotlivé iterace procesu však mohou dosahovat vyšších hodnot ztrátové funkce. Tato metoda učení je často vysvětlována na principu hraní počítačové hry. Umělá neuronová síť se snaží naučit postup hrou tak, aby ve výsledku dosáhla konce hry. K takovému cíli však vede potenciálně více cest a zároveň existuje mnoho cest, které k cíli nevedou. Pokud během pokusu nedojde k cíli, je cesta vyhodnocena jako nevhodná, tudíž se jedná o zpětnou vazbu od prostředí. Takto se zkouší velké množství pokusů a zároveň se neuronová síť učí z pokusů předcházejících. Snahou je najít nejen správnou cestu k cíli, ale pokud je to možné, tak například také tu nejefektivnější. V případech využití této metody jsou většinou známa alespoň základní pravidla, která upřesňují, jak s daty nakládat. Po každé operaci, která je provedena, je vyhodnocen dopad na celkový stav prostředí a určena nákladová funkce pro daný krok. Cílem však není dosažení co nejnižší jednotlivé nákladové funkce, ale její výsledné kumulované hodnoty při dosažení výsledku. Jednotlivé nákladové funkce a ani pravidla učení nejsou většinou přesně známa, ale jedná se spíše o aktuální odhady. Využití této metody učení je rozšířeno právě ve zmiňovaných počítačových hrách, nebo například při řešení sekvenčních úkolů.

Jednotlivé metody se navíc pro kombinaci jejich výhod a nevýhod mohou různě kombinovat. Například zpětnovazební učení spolu s učením s učitelem. V takových algoritmech je alespoň část založena na učení pod dohledem, s předem známým výsledkem, zatímco část může vycházet z poznatků natěžených v této fázi k učení na dalším setu dat. Algoritmus se například také může vydat cestou učení zcela nových dat, nebo cestou rozvíjení již získaných znalostí.

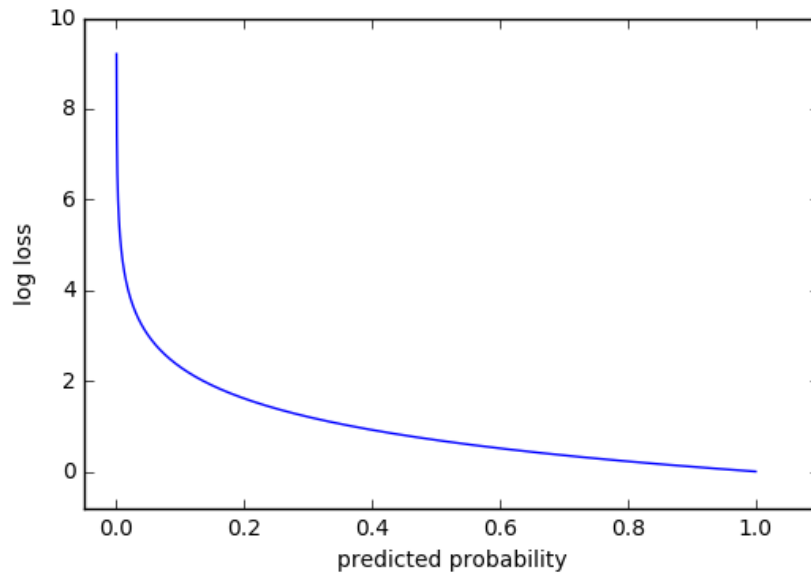


2.1.1 Chybové funkce

Střední kvadratická chyba - Mean Squared Error (MSE) je funkce vyjadřující sumu druhých mocnin rozdílů očekávaného a skutečného výsledku. Díky tomu, že počítá s existujícím a očekávaným výsledkem, využívá se jako indikátor během učení s učitelem. Tato funkce je vyjádřena pomocí rovnice:

$$MSE = \sum_{i=0}^n (y_i - h(x_i))^2, \quad (2.1)$$

kde y_i je hodnota z vektoru predikcí, n je velikost tohoto vektoru, a $h(x_i)$ je hodnota z vektoru očekávaných výsledků.



Obrázek 2.3: Závislost hodnoty křížové entropie na predikované pravděpodobnosti správnosti výsledku. [17]

Křížová entropie – Cross Entropy (CE) je funkce, která popisuje výkonnost modelu o více výstupních třídách, které se vzájemně nevyklučují. Popisuje rozdílnou distribuci pravděpodobnosti, se kterou daný objekt patří do určité výstupní třídy. Na obrázku 2.3



je znázorněna logaritmická funkce, která popisuje závislost hodnoty křížové entropie na předpovídané pravděpodobnosti správnosti určení výsledku. Samotná křížová entropie pro dvě výstupní třídy, což je nejčastější použití, je pak popsána rovnicí:

$$CE = \sum_{x \in \mathcal{X}} p(x) \log q(x), \quad (2.2)$$

kde $p(x)$ vyjadřuje pravděpodobnost, že výsledek náleží do třídy p , a $q(x)$ je pravděpodobnost, že výsledek náleží do třídy q . Nosič funkce \mathcal{X} je taková část definičního oboru, ve kterém funkce nabývají nenulovou hodnotu.

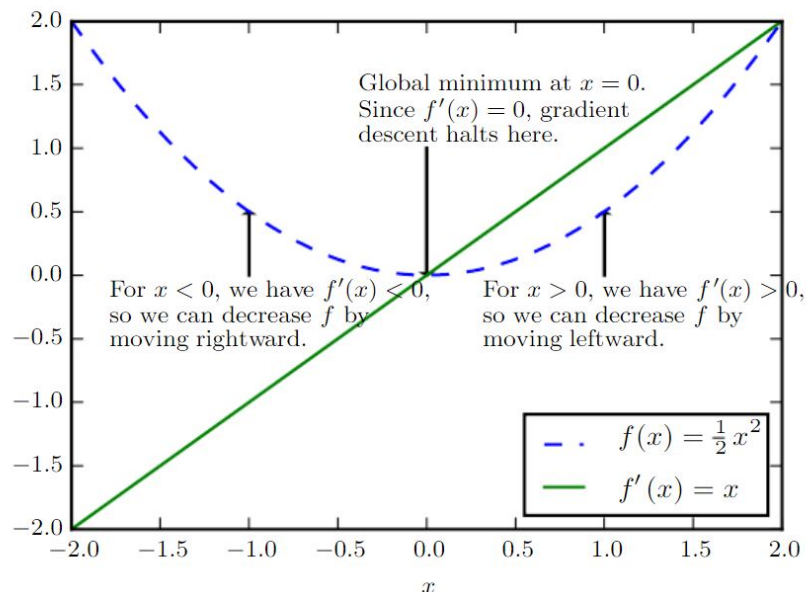
Gradient sestupu - Gradient descent (GD) je funkce nebo optimalizační algoritmus, který se využívá ke hledání lokálního minima funkce. Principem je iterační postup, kdy se postupuje po jednotlivých malých krocích proti směru gradientu funkce, nebo jeho přibližné hodnoty, v daném bodě funkce. Gradient funkce je totiž nejstrmější směr k jejímu lokálnímu maximu, proto se v případě gradientu sestupu postupuje proti směru gradientu funkce, jinak také známému jako gradient vzestupu. Princip funkce je znázorněn na obrázku 2.4. Jedná se o funkci prvního řádu, proto nabývá gradient sestupu hodnotu $f'(x) = 0$ ve svém minimu. Výpočet změny gradientu sestupu v umělých neuronech se provádí pomocí rovnice:

$$w_{i+1} = w_i - \eta \nabla G(w_i), \quad (2.3)$$

kde w vyjadřuje hodnotu z váhového vektoru, η vyjadřuje velikost kroku učení, udávající rychlost změny vah, a ∇G vyjadřuje aktuální velikost gradientu v dané iteraci. Vzhledem k tomu, že gradient vyjadřuje směr k maximu funkce, je v této rovnici hodnota gradientu odečítána.

Stochastický gradient sestupu – Stochastic gradient descent (SGD) je principiálně velmi podobný gradientu sestupu. Rozdíl je však v tom, že gradient sestupu je počítán z celého objemu dat, zatímco stochastický gradient sestupu je určován pouze z vybraného setu dat,

nikoliv z celého datasetu. To je obzvláště výhodné především při práci s velkými objemy dat, kdy práce pouze s podmnožinou dat značně snižuje výpočetní nároky celého algoritmu. Díky tomu se jedná o velmi důležitou metodu pro optimalizaci zpracování dat během strojového učení umělých neuronových sítí.



Obrázek 2.4: Gradient sestupu je obrácená hodnota první derivace funkce $F(x)$. [18]

2.1.2 Inicializace vah

Inicializace vah umělých neuronů je podstatnou součástí procesu učení. V této fázi, která předchází samotnému trénování neuronové sítě na danou problematiku, je jednotlivým neuronům přiřazena určitá hodnota váhy. Dohromady tyto váhy pro všechny neurony tvoří váhový vektor. Na těchto hodnotách je velmi závislý samotný proces učení sítě, ať už z hlediska rychlosti učení, tak například výsledné přesnosti. Pokud jsou navíc váhy špatně nastaveny u velmi komplikovaných neuronových sítí s více než pěti vrstvami, stává se jejich učení velmi obtížné.



Pro prvotní nastavení vah neuronů se proto velmi často používají hodnoty ze sítě, která již byla použita s podobným typem dat. První vrstvy v neuronových sítích obecně rozpoznávají základní tvary, je proto možné váhy z jiných sítí použít, pokud již trénovaná data typově odpovídají datům, která se chystáme nově učit neuronovou sítí. Tato metoda prvotního nastavení vah se nazývá *fine-tuning*, nebo *transfer learning*.

2.1.3 Zpětné šíření chyby

Zpětné šíření chyby je jeden z nejvyužívanějších algoritmů, které jsou využívány při učení neuronových sítí. Využívá se až v 80 % případů všech aplikací neuronových sítí, především pak u sítí s dopředným postupem učení s učitelem. Informace se v těchto sítích šíří pouze ve směru kupředu, takže nedochází k přijímání zpětné vazby od následujících vrstev neuronové sítě. Informace přichází na vstupní vrstvu sítě, následně postupně prochází skrytými vrstvami, a nakonec přichází do vrstvy poslední, výstupní. Funkce zpětného šíření chyby je speciální forma funkce gradientu sestupu. Pomocí této metody lze spočítat, jakým podílem se jednotlivé neurony podílejí na celkové chybě výsledku a na základě toho upravit jejich váhy.

Celý proces zpětného šíření chyby se skládá ze tří částí, kdy nejprve musí vstupní informace projít skrz celou neuronovou sítí. Na jejím konci je výsledek porovnán s očekávanou hodnotou a je určena chyba. Ve druhé fázi je tato chyba zpětně šířena sítí od konce na její začátek a při tom jsou určeny podíly na kumulované chybě. V posledním kroku dochází k úpravě váhového vektoru podle dříve určených parametrů. [zdroj 34] Přesnost určení chyby a tím pádem i celého procesu učení je silně závislá na velikosti kroku učení. Čím větším krokem se postupuje, tím větší je na konci kumulovaná chyba, a to může nepříznivě ovlivnit přesnost určení podílu chyby jednotlivých neuronů. Při příliš malém kroku je sice přesnost vyšší, ale celý algoritmus trpí vysokou výpočetní náročností a učení je mnohem delší. Vždy je proto potřeba hledat určitý kompromis mezi přesností a rychlostí učení.



2.2 Aktivační funkce

Aktivační funkce v umělém neuronu zastává jeho výpočetní funkci. Do neuronu může vstupovat několik hodnot zároveň, každá z těchto hodnot může mít jinou váhu, neuron může mít stanovenou i hodnotu bias, o kterou jsou vstupní hodnoty upraveny. Tato hodnota bias je předem známá a přičítá s sumě vstupů před aktivační funkcí. Každý neuron má však pouze jeden výstup, proto obsahuje aktivační funkci, která všechny vstupy a proměnné zpracuje podle daného postupu a vytvoří jeden výstup. Neuron nemusí mít na svém výstupu pouze binární hodnoty, jako je tomu u klasických jednodušších výpočetních metod. Volba vhodné aktivační funkce pak významně ovlivňuje rychlost výpočtů při učení neuronových sítí. Aktivační funkce je obecně definována rovnicí:

$$a = \sigma(w * x + b), \quad (2.4)$$

kde a vyjadřuje výstupní hodnotu z aktivační funkce, tedy i neuronu, σ je zástupce jakékoliv aktivační funkce, w je hodnota váhy daného vstupu, x vyjadřuje samotný vstup do funkce a b je hodnota bias.

Prahová funkce

Prahová funkce, jinak také nazývaná jako jednotkový skok, je nejjednodušší aktivační funkcí, se kterou se můžeme setkat. Funkce je závislá pouze na prahové hodnotě. Pokud je hodnota součtu vstupů, násobených jim odpovídajícími vahami vyšší než daná prahová hodnota, je výsledek této aktivační funkce $f(x) = 1$. Pokud je tento součet nižší než prahová hodnota, je výsledkem této aktivační funkce $f(x) = 0$. Prahovou funkci tedy můžeme definovat následujícím způsobem:



$$f(x) = \begin{cases} 1 & \text{pro } x \geq t \\ 0 & \text{pro } x < t \end{cases}, \quad (2.5)$$

kde x je vstupní hodnota do aktivační funkce a t je prahová hodnota dané aktivační funkce. Na obrázku 2.5b je grafické znázornění jednotkového skoku.

Lineární funkce

Lineární funkce, jinak také nazývaná přenosová funkce, nebo identické zobrazení. Funkce identity přiřazuje k prvku množiny úplně stejný prvek, takže se její hodnota nemění. Výstup je tedy proporciálně roven vstupům, konkrétně součtu všech vážených vstupních hodnot do neuronu, což je znázorněno na obrázku 2.5a. Funkce je definována základním vztahem:

$$f(x) = x, \quad (2.6)$$

kde x je součet všech vážených vstupů do aktivační funkce. Výhodou oproti prahové funkci je fakt, že nemá pouze binární výstup, který je nevhodný pro neuronové sítě. Její nevýhoda pak nastává při derivaci funkce podle x , protože hodnota derivace se nemění v průběhu celé funkce. Kvůli tomu nastává problém při využití algoritmu na výpočet gradientu sestupu, hledajícího minimum funkce. Další problém, který s sebou tato funkce nese, souvisí s přímou závislostí vstupů a výstupů. Při využití ve vícevrstvých neuronových sítích je každá další vrstva aktivována výsledkem vrstvy předcházející. Díky lineární závislosti vstupů a výstupů je ale vstup do první vrstvy a výstup z vrstvy poslední pouhou lineární závislostí a všechny vrstvy by se tak daly nahradit pouze jednou vrstvou. Z těchto důvodů i tato aktivační funkce nenachází moc využití.



Logistická funkce

Logistická funkce je graficky znázorněna pomocí křivky ve tvaru S, tato funkce se také nazývá sigmoida a je znázorněna na obrázku 2.5c. Jedná se o funkci, která nemá binární výstup a nemá lineární závislost vstupů a výstupů. Funkce je obecně definována předpisem:

$$f(x) = \frac{L}{1 + e^{-k(x-x_0)}}, \quad (2.7)$$

kde x_0 je hodnota funkce ve středu sigmoidu, L je maximální hodnota, kterou funkce nabývá, a k vyjadřuje rychlost růstu hodnoty funkce od jejího středu dále. Vzhledem k tomu, že proměnné L a k nabývají standartně hodnot 1 a hodnota x_0 nabývá hodnoty 0, používá se častěji spíš předpis:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2.8)$$

Logistická funkce může nabývat hodnot (0,1), proto se často využívá pro algoritmy určující pravděpodobnost. Nevýhodou této aktivační funkce je fakt, že v bodech vzdalujících se na ose x od středu se derivace funkce blíží nule, proto je zde velmi pomalý postup při využití gradientu sestupu, a to výrazně zhoršuje rychlost učení sítě. Tento fenomén se nazývá vanishing gradient. Tento problém se dá řešit vyšším výpočetním výkonem, nebo využitím jiné metody výpočtu.

Hyperbolický tangens

Hyperbolický tangens (\tanh) svým tvarem velmi připomíná logistickou funkci, liší se však ve strmosti stoupání v okolí středu funkce a hyperbolický tangens může nabývat hodnoty(-1,1). Stejně jako logistická funkce, tak i hyperbolický tangens má problémy s tzv. vanishing gradient. V okolí nuly má funkce vyšší gradient než sigmoida. Na rozdíl od ní je však tato



funkce centrována na hodnotě 0, díky čemuž je jednodušší její aplikace. Výhodou funkce je fakt, že negativní vstupy jsou funkcí určeny jako silně negativní, díky jejímu přesahu do záporných hodnot. Kvůli většímu gradientu v okolí nuly ale funkce postrádá přesnost v tomto regionu oproti funkci sigmoidy. Jak funkce *tanh*, tak sigmoida jsou využívány v dopředných neuronových sítích, hyperbolický tangens se pak často využívá při klasifikaci objektů do dvou tříd. Funkce je definována následující rovnicí:

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}, \quad (2.9)$$

kde x je hodnota vstupu do aktivační funkce. Grafické znázornění funkce hyperbolického tangentu se nachází na obrázku 2.5d.

ReLU

ReLU (Rectified Linear Unit) je aktivační funkce, která se dělí na dvě části. Pro záporné vstupy je výstupem funkce 0 a pro kladné vstupy se ReLU chová jako lineární funkce. Grafické znázornění této funkce je na obrázku 2.5e. Do roku 2011 byly široce využívány především funkce *tanh* a sigmoida, ale oproti nim má Relu značnou výhodu ve vypořádání se se zápornými vstupy. Pokud totiž do aktivační funkce přijde záporný vstup, funkce se vůbec neaktivuje a tím se značně šetří výpočetní výkon. Zatímco dvě zmíněné funkce jsou založeny více na fungování biologických neuronů, funkce ReLU se zaměřuje především na optimalizaci výpočtů v umělých neuronových sítích. V roce 2011 bylo zjištěno, že s pomocí této aktivační funkce lze dosáhnout lepších výsledků učení v hlubokých neuronových sítích. Funkce ReLU je definována předpisem:

$$f(x) = \begin{cases} 0 & \text{pro } x < 0 \\ x & \text{pro } x \geq 0 \end{cases}, \quad (2.10)$$



kde x je hodnota vstupu do aktivační funkce. Výhodou funkce je fakt, že při statistickém rozdělení dat se celkem 50 % neuronů vůbec neaktivuje a tím se snižuje náročnost výpočtu. Další výhodou je minimalizace fenoménu vanishing gradient, se kterým se potýká jak sigmoida, tak tanh. Problémem pro využití funkce může být to, že není ohraničena žádným intervalem, nedá se určit její střed, a v neuronových sítích se můžou vyskytnout ReLU neurony, které se dostanou do neaktivního stavu a nereagují na žádné vstupy. Tento stav je určitou formou vanishing gradientu a nastává hlavně při velkém tempu učení sítě, což pak snižuje jeho kapacitu učení. Problém se dá částečně řešit pomocí úpravy funkce podle následujícího předpisu:

$$f(x) = \begin{cases} 0.01x & \text{pro } x < 0 \\ x & \text{pro } x \geq 0 \end{cases}, \quad (2.11)$$

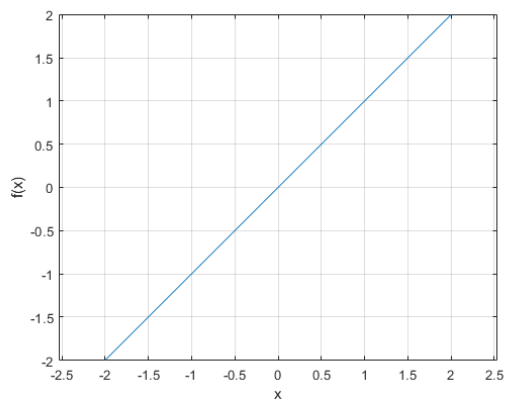
Tato funkce se nazývá Leaky ReLU a potlačuje problém neaktivních neuronů, ovšem za cenu mírného snížení kapacity sítě. Tato aktivační funkce je však jednou z nejčastěji používaných v neuronových sítích obecně.

Gaussova funkce

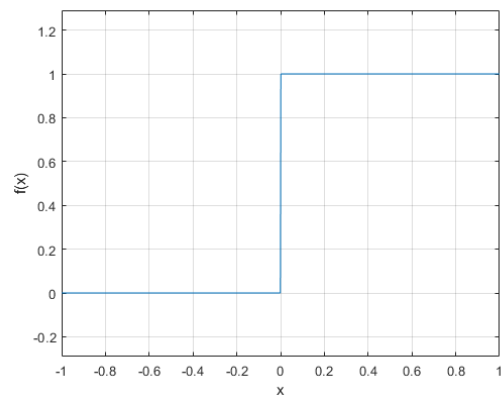
Gaussova aktivační funkce se nejčastěji v neuronových sítích využívá ve statistice s cílem popsat normální rozložení dat v datasetu. Dále se může využívat například při automatizovaném zpracování signálu na bázi umělé inteligence, kde se využívají gaussovy filtry pro redukci šumu. Tato funkce se dá využívat i při zpracování obrazu pomocí neuronových sítí, a to především k rozmazávání ostrých přechodů, smazávání nedokonalostí apod. Funkce je popsána rovnicí:

$$f(x) = e^{-x^2}, \quad (2.12)$$

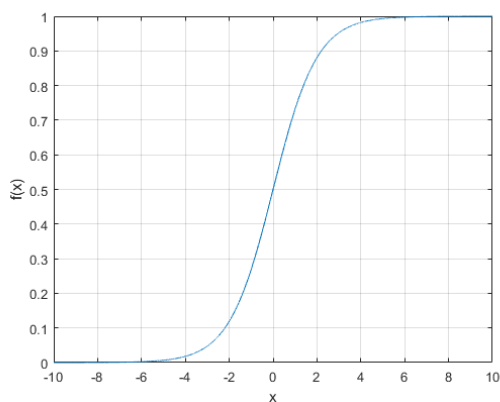
kde x je hodnota vstupu do funkce. Gaussova křivka je pak zobrazena na obrázku 2.5f.



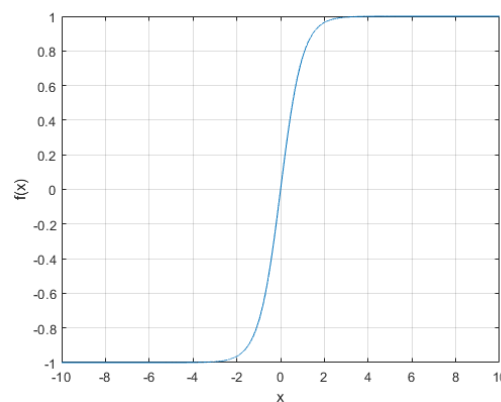
(a)



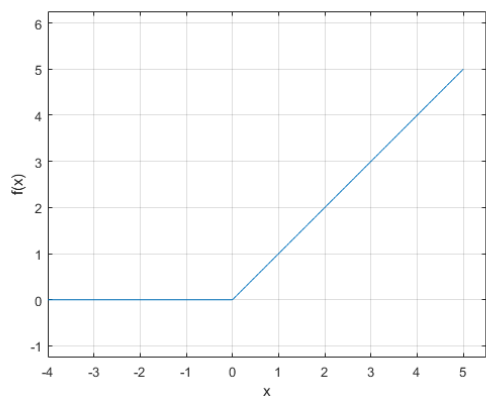
(b)



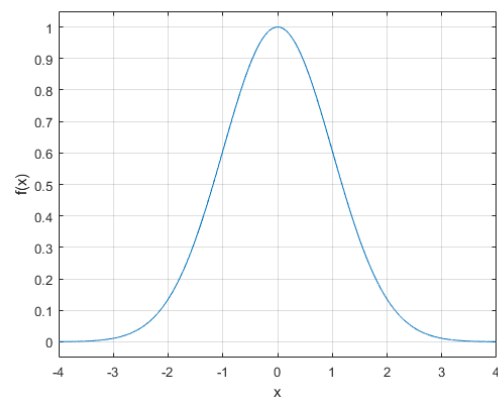
(c)



(d)



(e)



(f)

Obrázek 2.5: Grafické znázornění aktivačních funkcí umělých neuronových sítí. Na obrázku (a) lineární funkce, (b) prahová funkce, (c) logistická funkce, (d) hyperbolický tangens, (e) ReLU, (f) Gaussova křivka.



2.3 Vícevrstvé neuronové sítě

Vícevrstvé neuronové sítě jsou klasifikovány jako ty, které obsahují dvě a více vrstev neuronů. První vrstva přijímá vstupní data, která posílá dále do hlubších vrstev. Vrstvy, které se nachází mezi vstupní a koncovou vrstvou se nazývají skryté vrstvy. Pokud Neuronová síť obsahuje více než jednu skrytou vrstvu, nazýváme takovou síť hluboká neuronová síť (Deep Neural Network). Sítě, které se využívají pro zpracování obrazu jsou založeny především na konvolučních vrstvách, ale obsahují i plně propojené, pooling, nebo dropout vrstvy.

Plně propojené vrstvy

Plně propojená vrstva, anglicky dense layer, je složena z neuronů, kdy každý neuron přijímá výstupní hodnoty ze všech neuronů předchozí vrstvy nebo ze vstupního vektoru. Pokud je síť tvořena pouze plně propojenými vrstvami, nazývá se plně propojená neuronová síť. Takto navržené sítě mají vysokou schopnost učení a dosahují vysoké přesnosti určení výsledku, jejich plné propojení ale enormně zvyšuje nároky na výpočetní výkon. Každý pixel obrázku musí mít jeden odpovídající neuron, ale každý neuron přijímá data od všech pixelů. To znamená, že při zpracování obrázku o rozměrech 100 x 100 px má každý vektor vstupů a vektor vah 10000 prvků [19]. Proto se plně propojené neuronové sítě nevyužívají a plně propojené neuronové vrstvy se používají nejčastěji jako poslední, výstupní vrstva sítě. Tam se dobře hodí na finální klasifikaci oblastí zájmu a objektů.

Konvoluční vrstva

Konvoluční vrstva je nejčastěji využívanou vrstvou pro aplikace na klasifikaci objektů, ale i přes to vyžaduje kombinaci s dalšími vrstvami pro vytvoření použitelné neuronové sítě. Vzhledem k tomu, jak jsou plně propojené vrstvy náročné na počet vstupů i při zpracování velmi malých obrázků, využívají se vrstvy konvoluční. Ty nezpracovávají obrázek jako celek,



ale soustředí se na menší výřezy původního obrázku. Místo obrázku o velikosti 100 x 100 px se tak vrstva soustředí pouze na výřez o rozměru například 3 x 3 px. To místo celkového počtu 10000 vstupů redukuje vstupy pouze na 9 (při využití RGB jde o 27 vstupů). Takový čtverec se pak postupně aplikuje na celý obrázek. Posun čtverce po obrázku je udáván délkou kroku, tzv. stride, a jednotlivá okna se v průběhu posunu navzájem překrývají. Vektor vah je na všech pozicích sdílený, díky čemuž jsou sníženy nároky na výkon. Výsledkem konvoluční vrstvy jsou data o stejném rozměru jako měl vstupní obrázek. Data se nazývají aktivační mapa, nebo feature map a mají formu dvoudimenzionální mapy, na které je zobrazeno, kde došlo k aktivaci konvolučního jádra. Tato jednotlivá jádra, jinak nazývaná jako konvoluční filtry mají daný set parametrů, tedy vektoru vah, který je v průběhu kroku neměnný. Učení probíhá optimalizací parametrů těchto filtrů, kdy při tomto procesu dochází k postupné změně hodnot vektorů vah. Obecně je matematická operace konvoluce popsána rovnicí:

$$f(t) = (x * w)(t) = \int x(a)w(t - a)da \quad (2.13)$$

kde $x(t)$ je vstup funkce a $w(t)$ znázorňuje konvoluční jádro, tedy hodnotu z vektoru vah. Vzhledem k tomu, že data přicházející do konvolučních vrstev jsou převážně diskrétní, využívá se diskrétní konvoluce. Ta je definována následující rovnicí:

$$f(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t - a) \quad (2.14)$$

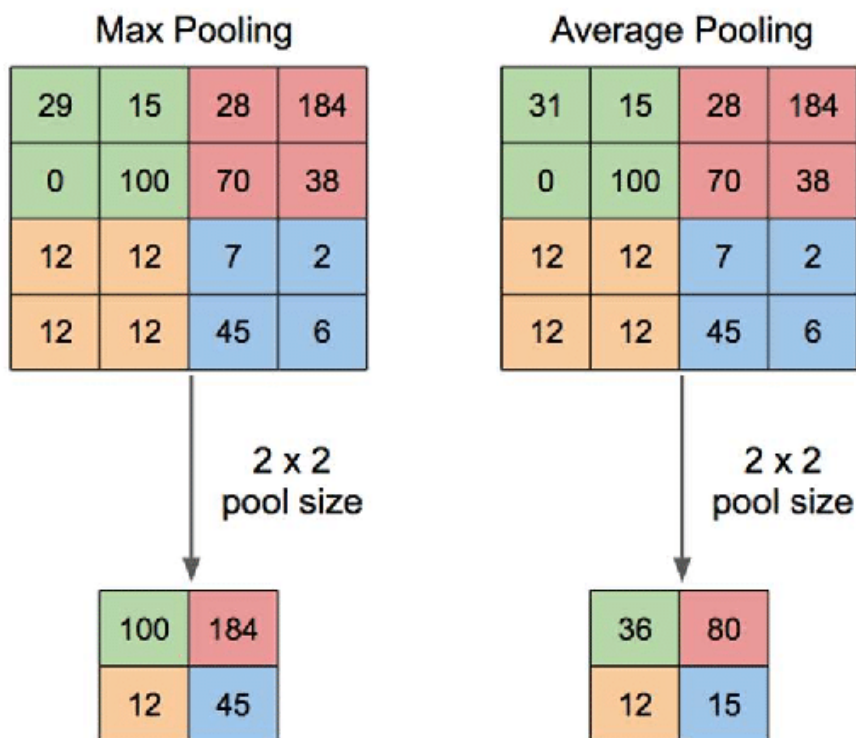
kde $x(t)$ je vstup funkce a $w(t)$ znázorňuje konvoluční jádro, tedy hodnotu z vektoru vah.

Pooling vrstva

Pooling vrstva slouží v neuronové síti k redukci neuronů v následující vrstvě. V pooling vrstvě totiž dochází ke spojení výstupů několika neuronů do jednoho neuronu v další vrstvě. Tím se tak efektivně sníží počet neuronů a s tím i výpočetní náročnost operací. Pooling



vrstva většinou následuje po konvoluční vrstvě. Při sdružování několika neuronů do jednoho, tzv. clusteringu, je více způsobů, jak určit výslednou hodnotu, kterou má přebrat následující neuron. Nejčastěji je využíván max pooling, kdy hodnota následujícího neuronu je určena jako maximální hodnota z daného clusteru. Tento princip je znázorněn na obrázku 2.6. Proti maximální hodnotě se může vyskytovat také opak, min pooling, kdy výsledná hodnota je rovna minimu z daného clusteru. Další možností, která se využívá je average pooling, kdy se výsledná hodnota rovná průměrné hodnotě všech výstupů neuronů v daném lokálním clusteru. Tento způsob je také znázorněn na obrázku 2.6. Agregaci můžeme využívat lokální, kdy probíhá pooling pouze na malém počtu neuronů, nebo můžeme mít agregaci globální, kdy se slučují mnohem větší oblasti neuronů.



Obrázek 2.6: Znázornění principu funkcí max pooling a average pooling. [20]



Dropout vrstva

Tato vrstva slouží především k potlačení tzv. přeučení neuronové sítě. V takovém případě dochází k tomu, že síť se až moc perfektně naučí na daném vzorku dat, že pak dále není schopná pružně reagovat na další příchozí data, nebo provádět další výpočty spolehlivě. V každé fázi učení proto určitý počet neuronů automaticky nastaví na výstupní hodnotu 0. Tím sice dochází ke snížení kapacity sítě, ale jednotlivé neurony se nemohou spoléhat na nepřerušované vstupy od okolních neuronů a jsou nuceny se učit. K této deaktivaci dochází neustále při fázi učení, do výsledků tak proniká více šumu. To ovšem síť dělá ve výsledku stabilnější a odolnější vůči podobným vnějším vlivům [21].

2.4 Detekce objektů

První způsoby detekce oblastí zájmu spočívaly především ve srovnávání barvy a tvaru objektů, což bohužel neumožňuje dosažení uspokojivých výsledků. Přelom nastal především s novým tisíciletím, kdy začal skokově narůstat výpočetní výkon a tím i možnosti pro složitější algoritmy pro detekci objektů. Tyto detektory však byly většinou založeny na principu dělení scény na velký počet oblastí, které byly jednotlivě vyhodnocovány. Pro lepší výsledky navíc musely být vyhodnocovány opakovaně s použitím různých vstupních parametrů, což i s rostoucím výkonem pořád značně limitovalo jejich využití. Zásadní zlom pak nastal především s nástupem a rozvojem umělých neuronových sítí, tedy základu v oblasti umělé inteligence. Tyto umělé neuronové sítě svou architekturou napodobují neurony živých organismů. Tyto umělé sítě se dokážou učit, trénovat na širokých databázích dat a následně odhalit vzorce a rozpoznávat objekty s vysokou úspěšností. [20] Princip detektorů spočívá právě v tom, že se snaží v obrazových datech hledat určité vzorce a specifické vlastnosti, které od sebe pomáhají jednotlivé objekty odlišovat. Detektory navíc nemusí být založeny pouze na neuronových sítích, ale i na klasických výpočetních metodách. Rozdíl je však v tom, že nejprve musí být přesně definovány jednotlivé objekty a popsány vzorce a vlastnosti, na



základě kterých se budou objekty klasifikovat. Takový systém ale nedokáže pružně reagovat na změnu podmínek a učit se například změny, které během detekce nastávají. Neuronové sítě disponují vlastností, že jsou schopny se průběžně adaptovat na podmínky a učit se na základě příchozích dat. Typicky se pak jedná o konvoluční neuronové sítě. Nejčastější detektory na základě neuronových sítí jsou R-CNN (Region-based Convolutional Neural Networks) a jeho možné varianty, SSD (Single Shot MultiBox Detector) a YOLO (You Only Look Once).

Výsledná úspěšnost zvoleného detektoru objektů závisí především na typu využití neuronové sítě. Ty se od sebe velmi zásadně liší, zatímco některé architektury obsahují například 11 vrstev, některé dosahují i rozměrů 152 vrstev, což z takové sítě dělá jednu z nejrozsáhlejších vůbec. Jednou z nejčastěji využívaných neuronových sítí je AlexNet od týmu Alexandra Krizhevského. Ti s touto sítí v roce 2012 vyhráli na mezinárodní soutěži neuronových sítí s chybovostí určení objektu 15,3 %. Tato síť využívá detektor objektů R-CNN.

Další neuronovou sítí, která principiálně vychází z AlexNet, je model ZF Net. Tento model obsahuje ve svých vrstvách převážně aktivační funkci ReLU, jako chybovou funkci využívá křížovou entropii, a dosahuje průměrné chybovosti 11 %. Další sítě vznikají každým rokem ve snaze zlepšit jejich výsledky při vynaložení co nejmenšího výpočetního výkonu. Co se přesnosti týče, na samotném vrcholu se nachází neuronová síť s názvem ResNet od výzkumného týmu společnosti Microsoft. Tato síť se skládá ze 152 neuronových vrstev a dosahuje průměrné chybovosti 3,5 % [22]. Tato hodnota je již na samé hranici možností, zbývá velmi malý prostor pro zlepšení přesnosti, která v mnoha případech převyšuje i hodnoty chybovosti dobře trénovaného člověka. Další sítě se proto mnohem častěji ubírají směrem snížení výpočetní náročnosti a zvýšení rychlosti řešení.

R-CNN

Region-based Convolutional Neural Network, a především jeho modernější modifikace, jsou nejčastěji používané metody pro spolehlivou detekci objektů. Detektor je rozdělen na dvě



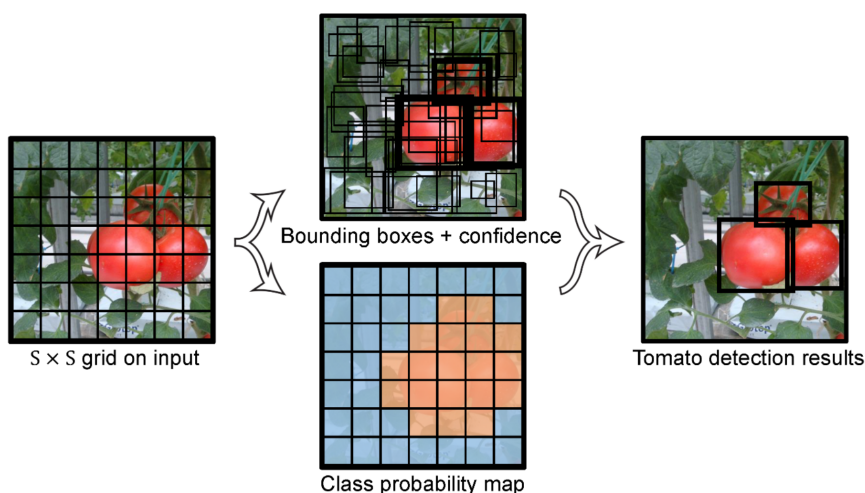
části, kdy v první fázi model nejprve vyhledává potenciální oblasti, kde by se mohly detekované objekty nacházet a následně probíhá analýza těchto oblastí s cílem zjistit, jestli opravdu obsahují objekty, které se snažíme detekovat. Detektor těchto oblastí vytvoří většinou okolo 2000 za použití speciálního algoritmu. Ten porovnává blízké pixely a pokud jsou si barevně podobné, předpokládá, že by se v této oblasti mohl objekt nacházet. Vybrané oblasti následně prohází konvolučními vrstvami a v poslední vrstvě jsou klasifikovány. Tato vrstva je založena na metodě podpůrných vektorů, což je metoda strojového učení s učitelem. Základní verze detektoru R-CNN dosahuje průměrně zhruba 30% úspěšnost určení objektů.

Detektor byl v následujících letech několikrát vylepšen na verze Fast R-CNN a Faster R-CNN. Tyto detektory již nevyužívají dvoufázový postup jako jejich starší předchůdce, navíc poslední vrstva založená na metodě podpůrných vektorů byla nahrazena jednou softmax vrstvou. Softmax je jeden z typů aktivační funkce v neuronu, tato vrstva je pak složena právě z těchto neuronů. Faster R-CNN kombinuje trochu jiné vrstvy, hlavní rozdíl je ale ve tvoření oblastí s potenciálním výskytem objektů. Zatímco R-CNN potřebuje v první fázi zhruba 2000 oblastí, Faster R-CNN stačí pouze 300. Díky tomu se celý proces detekce podstatně zrychlil a dovoluje zpracování zhruba 5 snímků za sekundu [23].

YOLO

YOLO – You only look once je detektor, který na rozdíl od většiny ostatních detektorů provádí ohraničení oblastí a následnou detekci v jedné fázi. To znamená, že data od vstupu projdou neuronovou sítí pouze jednou a ta rovnou poskytne výsledek. Pomocí jedné neuronové sítě tak tento detektor dokáže vyhledat oblasti s potenciálním výskytem objektů, ale zároveň je dokáže i klasifikovat. Tato vlastnost detektoru zajišťuje jeho nižší výpočetní nároky oproti dvoufázovým detektorům. Výsledkem detekce je pravděpodobnost, se kterou se hledaný objekt nachází v určitém ohraničeném prostoru.

Detektor kombinuje tři různé techniky detekce, které ve výsledku tvoří velmi robustní algoritmus schopný přesné detekce a schopný se velmi dobře učit. První metoda obrázek rozdělí do pravidelné mřížky. Jednotlivé čtverce v mřížce mají identické rozměry a algoritmus určuje, jaký objekt se v nich nachází a zároveň definují pravděpodobnost, s jakou se zde objekt vyskytuje. Tím vznikne mapa pravděpodobnosti výskytu jednotlivých tříd objektů definována na vytvořené mřížce. Další metoda rovnou ohraničuje objekty v obraze. Během této fáze dochází ke klasifikaci objektů, určení jejich pozice a pravděpodobnosti správnosti určení. Poslední metoda porovnává odhadované ohraničení objektu s jeho reálnými hranicemi a cílem je zajistit, aby se tyto dvě ohraničení shodovala. Kombinací všech těchto metod, především ale kombinací metody s klasifikační mřížkou a metody ohraničování objektů pak dosáhneme přesných výsledku s pomocí metody YOLO. Princip kombinace těchto metod je znázorněn na obrázku 2.7.



Obrázek 2.7: Princip kombinace jednotlivých metod, dohromady tvořící detektor YOLO. [24]

2.5 Parciální závěr

Ústav letecké dopravy fakulty dopravní ČVUT má k dispozici zařízení pro sledování pohybu očí při využívání brýlí pro rozšířenou realitu. V principu se jedná o kamery, které jsou umístěny na brýlích a snímají okolní scénu a obě oči. Proto je zde možnost navrhnout řešení,



kteřé tyto kamery umožní nainstalovat na jiný nosič a využívat je při aplikacích a interakci s reálným prostředím.

Cílem praktické části je návrh právě takového řešení, které bude pohodlné i při dlouhodobém měření, nebude výrazně omezovat zorné pole člověka a bude jednoduché na výrobu prototypu. Na základě analýzy současného stavu týkající se především jednotlivých typů pohybů očí bylo vyhodnoceno, že dané zařízení svými parametry vyhovuje pro detekci pohybu očí, které jsou důležité při využití v kokpitu letadla, nebo podobných aplikacích. Jedná se především o pohyby s větší úhlovou výchylkou, které je možné spolehlivě detekovat i pomocí levnějšího vybavení. Není zde kladen takový důraz na detekci velmi jemných pohybů očí, které jsou vykonávány samovolně, bez přičinění vůle člověka. Dalším cílem této praktické části je vytvoření postupů pro zpracování získaných dat. Jedná se o software na prvotní zpracování naměřených dat a následně o software pro automatické rozpoznávání přístrojů a oblastí zájmu v kokpitu.

Následně experimentálně ověřím, že zařízení i po přestavbě na jiný nosič stále funguje přesně a podle předpokladu. Dále při experimentu získám velké množství dat, která následně poslouží pro práci se softwarem na detekci oblastí zájmů. Součástí práce bude vytvoření databáze objektů v kokpitu letounu Cessna 182, která může posloužit pro následná měření již bez potřeby další práce s tvorbou takové databáze.



3 Metody

Základní součástí navrhovaného systémového řešení je zařízení pro sledování pohybu očí. Toto zařízení je však původně určeno pouze k aplikaci s brýlemi pro rozšířenou realitu Epson BT-300. V práci se zaměřuji na využití těchto kamer i v reálném prostředí v kombinaci s běžně dostupnými slunečními brýlemi. Proto je nutné provést experimentální ověření, zda tento systém funguje podle předpokladů, a poskytuje relevantní a přesná data. Veškerá získaná data budou poté podrobena zpracování pomocí softwaru pro detekci oblastí zájmu. Primárním cílem je provedení experimentu a následná znovupoužitelnost v kokpitu letadla, systém však bude univerzální a využití může najít nejen v letectví nebo automobilovém průmyslu, ale například i při hodnocení reklam apod.

3.1 Návrh experimentu

Pro naplnění cíle práce, kterým je vytvoření uceleného systému pro záznam a následné vyhodnocení pohybu očí je třeba nejprve experimentálně ověřit samotné zařízení, které bylo vytvořeno jako prototyp. Umístění kamer přibližně odpovídá jejich umístění na brýlích pro rozšířenou realitu Epson BT-300, ale geometrie a přesné umístění se mohou lehce lišit. Z tohoto důvodu je tedy nutné ověřit, že zařízení je nadále funkční a změna geometrie neovlivňuje jeho přesnost. Celý proces ověřování je nahráván a údaje o pohybech očí jsou nejen vizuálně zobrazeny na monitoru, ale zároveň jsou data ukládána do souboru .csv.

Experiment, který v rámci práce provádím, probíhá v reálném prostředí pilotní kabiny letounu Cessna 182. Prvním krokem pro přesné měření je kalibrace. Software pro záznam dat o pohybech očí a záznam okolního prostředí má vestavěnou funkci kalibrace kamer. Před začátkem je proto potřeba provést tuto kalibraci. Po úspěšném provedení tohoto kroku následuje samotný experiment pro ověření funkčnosti zařízení. Nejprve se v pilotní kabině pořídí video záznam v délce 10-20 minut. Následuje pořízení několika víceminutových záznamů za letu. Toto je potřeba z důvodu dosažení různých podmínek v kokpitu, jako jsou



světelné podmínky, hodnoty na přístrojích, ale i jiné úhly pohledu v kokpitu a vliv vibrací. Během fáze měření na zemi je pak podstatné získání dat pro ověření přesnosti měření a následné vyhodnocení oblastí zájmu. Proto bude subjekt podle pokynů zaměřovat pozornost na jednotlivé předem určené přístroje. Výsledky automatického určení oblastí zájmu tak je možné porovnat se známými pozicemi přístrojů. Cílem je dosažení co nejvyšší přesnosti, k tomu vede především správné nastavení pozice kamer snímajících oči a dobře provedená kalibrace.

3.2 Zařízení

V této práci je využito zařízení, které se skládá z obrouček slunečních brýlí a soustavy kamer pro snímání pohybu očí a okolního prostředí. Toto řešení je navrženo s ohledem na možnost a materiál, kterým laboratoř v současnosti disponuje. Na trhu však existuje několik již hotových zařízení, která poskytují veškeré potřebné funkce a není potřeba je nijak upravovat a přizpůsobovat. Taková zařízení však stojí od 50 000 Kč až do 300 000 Kč v závislosti na kvalitě, značce a funkcích daného zařízení. Nejblíže k zařízení z této práce má systém Pupil Core od výrobce PupilLabs. Tento systém je vybaven stejnými kamerami, které jsou využívány v této práci, disponuje však pečlivě navrženými obroučkami pro co nejpohodlnější nošení. Takové zařízení se pohybuje v podobné cenové hladině, jako kamery, které využívám, tedy zhruba 50 000 Kč. Dalším dostupným zařízením je PupilLabs Invisible od stejné firmy. Důraz při návrhu tohoto zařízení byl kladen především na miniaturizaci kamer a jejich umístění do obrouček brýlí, jak je ukázáno na obrázku 3.1b.

Parametry kamer jsou opět stejné jako parametry těch, které jsou využity v této práci. Takové zařízení již však stojí okolo 140 000 Kč, využití proto nachází především v komerčním sektoru, nikoliv akademickém. Kromě funkce vzhledu totiž nepřináší žádnou zásadní výhodu. Jedním z největších výrobců eye trackingových zařízení je společnost Tobii. I tato společnost pak nabízí zařízení, které je svou funkcí podobné dříve zmíněným řešením, jejich zkušenosti a kvalita jsou však na vyšší úrovni. Kamery poskytují vyšší rozlišení i snímkovací frekvenci,



(a) Tobii Pro Glasses 3



(b) Pupil Labs Invisible

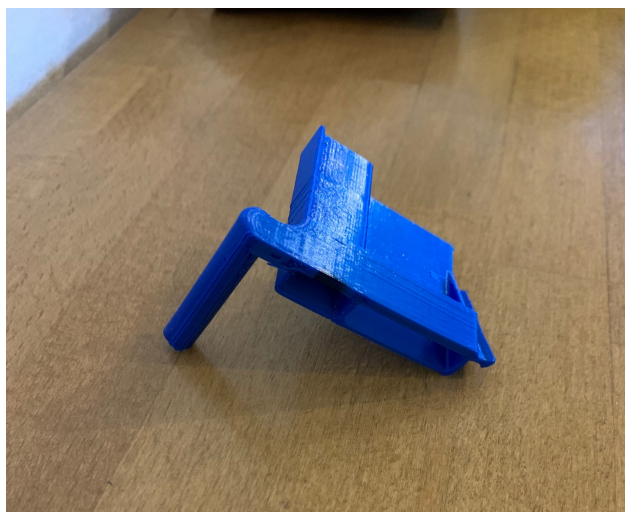
Obrázek 3.1: Aktuálně dostupná řešení pro detekci pohybu očí na trhu. Na obrázku (a) systém od společnosti Tobii, na obrázku (b) systém od společnosti Pupil Labs.

celkově mají větší počet kamer pro přesnější data, která jsou navíc méně náchylná na vliv okolních světelných podmínek. Na obrázku 3.1a je zařízení Tobii Pro Glasses 3, které je vzhledem velmi podobné PupilLabs Invisible. Díky vyššímu počtu kamer a jejich větší kvalitě je však podstatně dražší, cena se pohybuje nad 230 000 Kč. Jedná se však o řešení s nejlepšími vlastnostmi, proto pokud je potřeba extrémně přesných dat, která jsou navíc sbírána s velkou spolehlivostí, jde o jedny z nejčastěji používaných brýlí na trhu.

Pro vytvoření prototypu eye trackingových brýlí v mé práci bylo využito kamer, které byly původně určeny k brýlím pro rozšířenou realitu. Jedná se o systém od společnosti PupilLabs a skládá se z jedné kamery snímající okolní prostředí, a dvou kamer, kdy každá snímá jedno oko. Čelní kamera disponuje rozlišením 1280x720 pixelů a snímkovací frekvencí 30 Hz. Tyto hodnoty poskytují uspokojivou kvalitu záznamu při použití v kokpitu letadla, tedy maximálně do 1 metru vzdálenosti od přístrojů. Kamery pro snímání očí mají rozlišení 192x192 pixelů, pracují však při frekvenci 120 Hz. Tato vyšší hodnota snímkovací frekvence je podstatná pro detekci velmi rychlých pohybů očí, jako jsou sakády nebo mikrosakadické pohyby.

Vzhledem k tomu, že tyto kamery byly původně určeny pro jinou aplikaci, bylo potřeba vytvořit také vhodný upevňovací mechanismus na brýle, které jsou použity pro účely této práce. K tomu bylo využito návrhu vlastních držáků a jejich vytištění na 3D tiskárně. Při

návrhu se vycházelo z původního designu držáků kamer, které jsou určeny pro brýle pro rozšířenou realitu. Jejich úprava pro použití na obroučky slunečních brýlí však způsobovala nepohodlí při nošení a proto bylo potřeba držáky dále předělat. Na obrázku 3.2 je finální verze držáků kamer. Ty byly zeštíhleny, aby nebyly při nošení v kontaktu s hlavou, zároveň byly prodlouženy směrem dolů, což umožňuje umístění kamer podstatně níže. Tím se tak dostanou pod samotné obroučky brýlí a nemusí procházet skrz obroučky, jak tomu bylo u prvního prototypu.



Obrázek 3.2: Druhý prototyp držáku kamer vytvořeného s pomocí 3D tisku.

Samotné obroučky byly upraveny minimálně, byla z nich odstraněna sluneční skla, protože znemožňovala detekci pohybu očí. Na nožičkách brýlí pak byly pomocí lepicí pásky vytvořeny plochy, které způsobují zvýšené tření a těsný kontakt s držáky kamer. Tyto držáky se tak díky tomu nemohou samovolně pohybovat, což je podstatné pro samotné měření. Na obrázku 3.3 je finální podoba zařízení, které bylo vytvořeno pro účely této práce. Jedná se o druhý prototyp, který byl vytvořen.



Obrázek 3.3: Finální podoba brýlí pro sledování pohybu očí.

3.3 Příprava experimentu

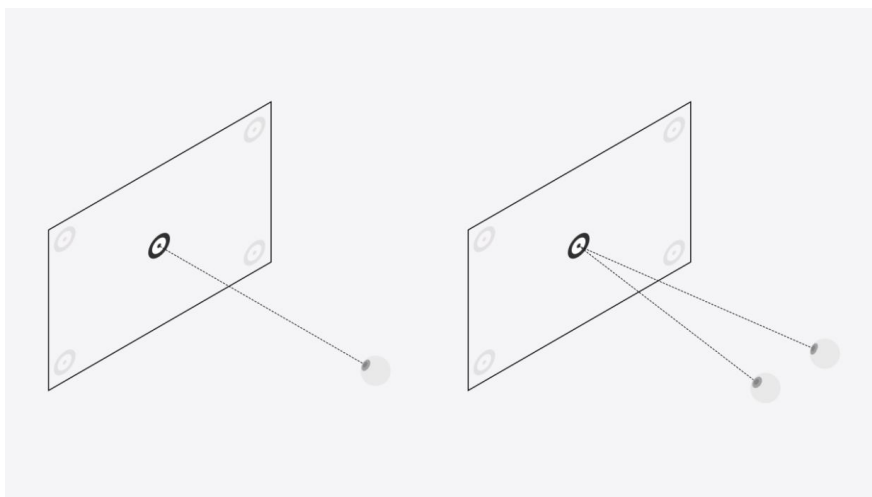
Kalibrace zařízení je základním předpokladem pro započítání měření. Před samotnou kalibrací je však potřeba provést několik úkonů, které dále zlepšují přesnost zařízení. Prvním krokem je pečlivé usazení brýlí na hlavě tak, aby bylo jejich nošení pohodlné a aby držáky kamer nebyly v kontaktu s hlavou. Hlavní snahou je dosáhnout co nejmenšího pohybu brýlí během kalibrace a měření. Drobný pohyb brýlí v průběhu času je bohužel nevyhnutelný a ovlivňuje přesnost dat, proto je nutné tyto pohyby redukovat na minimum. Dále je před začátkem kalibrace nutné nastavení polohy kamer snímajících oči. Tyto kamery detekují zornice, proto je důležité se ujistit, že jsou zornice v zorném poli kamer nejen při pohledu vpřed, ale i při maximálních výchylkách pohledu člověka. Dále je před kalibrací doporučeno mít brýle správně usazené a provést několik pohybů očí v různých směrech. To umožní programu vytvoření 3D modelu oka a dále zlepšit přesnost kalibrace a měření.



Po provedení těchto základních kroků je třeba provést kalibraci a validaci dat. Pro optimální výsledky je vhodné mít obrazovku s kalibračním obrazcem v podobné vzdálenosti, jako jsou následně vybrané oblasti zájmu samotného experimentu. Software pro záznam a zpracování dat sice dokáže hloubku scény odhadovat na základě úhlu, který vůči sobě svírají oči, může nám tak ale do měření vstupovat nepřesnost. Proto výrobce zařízení a softwaru doporučuje provádění kalibrace ve stejné vzdálenosti jako následné měření. Zároveň pro získání co nejpřesnějších dat je vhodné provádět měření na kratší vzdálenosti. Zařízení má podle způsobu vyhodnocování dat chybu měření 1-2°. Při pozorování oblasti zájmu ve vzdálenosti 1 metr a chybě měření 1° je odchylka 1,7 cm. Při vzdálenosti 2 metry a stejné chybě měření je tato odchylka již 3,4 cm a s rostoucí vzdáleností úměrně roste. Navíc s rostoucí vzdáleností oblasti zájmu se snižuje úhlová výchylka očí a tím se zmenšuje rozlišovací schopnost zařízení a přesnost dat. Jakmile se změna úhlu pohledu přiblíží chybě měření, ztrácí zařízení schopnost detekce těchto pohybů.

Samotná kalibrace zařízení má několik možných podob, které mohou být zvoleny podle aktuálního využití. Cílem kalibrace je dosažení korelace dat z kamery, která snímá zorné pole subjektu, a kamer, které snímají pohyby očí. Kalibrace podle aktuální konfigurace může probíhat jako monokulární nebo binokulární. Díky tomu, že zařízení, které je navrženo pro pokus v této práci, disponuje dvěma kamerami, je použita kalibrace binokulární. Proces kalibrace pak probíhá podle předem nastavených pravidel a scénářů, které se nazývají choreografie. Základní choreografií je zobrazování kalibračního terče na obrazovce počítače. Jedná se o základní choreografii pro rychlé uvedení zařízení do provozu. Tato choreografie má formu pěti po sobě následujících terčů, které je potřeba sledovat. Každý terč se zobrazí na několik sekund a je nutné mu věnovat maximální pozornost. Terče jsou na obrazovce postupně zobrazovány nejprve uprostřed a následně v každém z rohů, jak je znázorněno na obrázku 3.4 jak pro monokulární použití, tak pro binokulární použití.

Další choreografie najdou využití například v případě, kdy není možné provést kalibraci pomocí obrazovky. To umožňuje kalibrace založená na jediném terči, který může být zobrazen buď na monitoru, nebo vytištěn na papíře. Tento postup spočívá ve sledování tohoto



Obrázek 3.4: Základní choreografie pro kalibraci zařízení.

terče bez přerušení, zatímco hlavou vykonáváme pohyby po spirálovité trajektorii. Tímto se postupně oči dostanou do všech krajních poloh, zatímco sledují daný bod a systém může být zkalibrován. Posledním scénářem je kalibrace pomocí objektů nacházejících se v zorném poli subjektu. Tento proces spočívá ve sledování jakéhokoliv výrazného bodu v zorném poli, zatímco je tento bod na monitoru označen. Tím dojde k zaznamenání dat a s dostatečným počtem takto zaznamenaných bodů se dokáže systém kalibrovat. Minimální počet označených bodů je 9, s rostoucím počtem dochází ke zpřesňování hodnot odchylky měření. Tento postup je však doporučován pouze ve specifických situacích, kdy není možné použití žádné z předchozích dvou choreografií.

Po dokončení libovolné kalibrační choreografie je důrazně doporučováno provést ještě validaci dat. Ta spočívá opět v zobrazení několika bodů na monitoru a jejich pečlivému sledování. Program dokáže vyhodnotit odchylku od známých poloh a určí chybu přesnosti měření. Pokud je chyba příliš velká, je potřeba provést znovu proces kalibrace a následně validace. Až pokud z procesu validace vyjdou uspokojivé výsledky, je možné začít se samotným měřením.

Společnost PupilLabs zároveň doporučuje rozdělení měření na několik kratších úseků a mezi každým z těchto měření opakovat proces kalibrace a validace. Jen tak je možno dosáhnout maximální přesnosti výsledků. Při dlouhodobém měření nám do dat vstupují nepřesnosti



vzniklé pohybem brýlí na hlavě. Software dokáže pohyb brýlí na hlavě kompenzovat využitím výpočetního modelu založeném na 3D modelu oka, chyba měření se však může pohybovat v rozmezí 1-2°, zatímco při použití 2D modelu je chyba měření menší než 1°. Aplikace 2D modelu pro výpočet eye trackingových dat je doporučeno pro využití u experimentů, kde není třeba pohyb člověka a hlavy a kde může být měření rozděleno do několika kratších úseků. Cílem toho je dosažení co nejpřesnějších výsledků. Naopak využití 3D modelu je vhodné pro pokusy, kde je potřeba větší mobilita subjektu, a kde nejde rozdělit měření do několika menších bloků s průběžnou kalibrací zařízení.

Již při prvním nasazení brýlí je doporučeno zapnout nahrávání ze všech kamer. S tímto přístupem je nahráván i samotný proces kalibrace. Při následném zpracování záznamů je pak možné provést ze záznamu kalibraci post-hoc, v programu Pupil Player. Při této opakované kalibraci můžeme provést změny parametrů detekce pohybu očí, jako jsou minimální a maximální detekovatelné rozměry zornice. Zároveň pokud je měření prováděno v několika kratších blocích s průběžnou kalibrací, mohou být data z průběžné kalibrace použita k vyhodnocení akumulovaného posunu brýlí při předchozím měření. Pro optimální výsledky je také doporučeno se zařízení během měření nedotýkat, nicméně i po takovýchto dotycích na brýlích během měření nebylo zjištěno zhoršení přesnosti získaných dat.

3.4 Průběh experimentu

Experiment probíhal kompletně v kokpitu letounu Cessna 182 a byl rozdělen na dvě části. Posloupnost těchto dvou částí není důležitá, první bylo proto provedeno měření za letu. Podle doporučení bylo nahrávání zahájeno již před kalibrací zařízení. Pilotka měla brýle nasazené po dobu celého letu. Během měření nebylo možné záznamy rozdělit do několika kratších bloků a mezi nimi provádět novou kalibraci a validaci. Cílem měření za letu bylo dosažení realistických podmínek v kokpitu, různých světelných podmínek, ale také různých úhlů pohledu na jednotlivé přístroje. Let trval zhruba hodinu a skládal se ze dvou okruhů okolo letiště s přistáním a novým vzletem, respektive letným přistáním. Následoval delší úsek



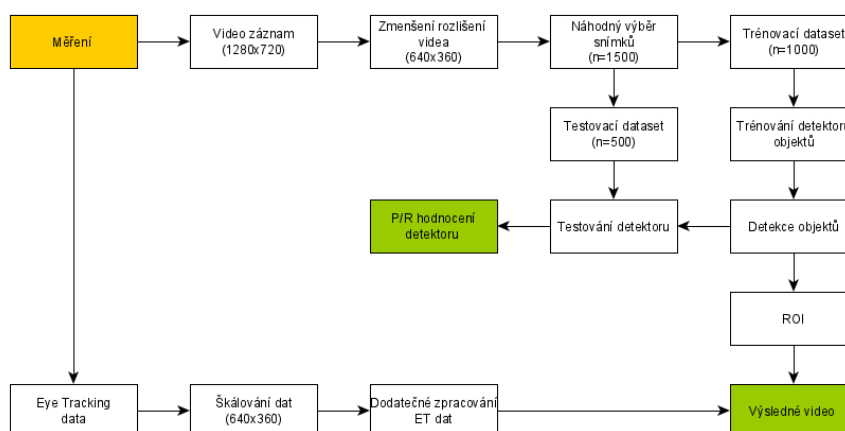
letové fáze, který trval přibližně 20 minut a následně poslední přistání. Během těchto činností je potřeba kontrola a interakce s palubními přístroji častější než při dlouhém a ustáleném letu, což bylo ideální pro potřeby měření. Při každém přistání bylo přerušeno nahrávání a ihned zahájeno nové, ovšem bez kalibraci mezi nimi. Toto dělení záznamů bylo prováděno především kvůli následné práci s nimi, aby byla redukována jejich velikost a náročnost na výpočetní výkon. Ve výsledku tak byly v této fázi pořízeny 3 záznamy, každý o délce přibližně 20 minut.

Druhou fází experimentu bylo snímání kokpitu na zemi, v ideálních podmínkách. Záznam z kokpitu má délku přes 10 minut a obsahuje pohledy na všechny přístroje na přístrojové desce letadla. Pohledy na přístroje jsou z různých úhlů, vzdáleností, nebo z trochu jiné pozice v kabině. Toto měření probíhalo za optimálních světelných podmínek, takže bez přímého světla do očí pilota, ale zároveň s dostatečným světlem v kokpitu. Pro nejlepší výsledek fáze měření na zemi bylo potřeba také zajistit minimální kontrast mezi tmavým přístrojovým štítem a světlým výhledem z kokpitu. Toho bylo dosaženo přistavením letadla do stínu, bez přímo osvětleného prostředí ve výhledu z kabiny. Před začátkem měření byla provedena několikanásobná kalibrace a validace, protože se zařízení z počátku nedařilo správně nastavit.

Součástí pozemní fáze bylo také jednoduché ověření kalibrace nezávisle na vestavěných procesech záznamového softwaru. Pilot se podle pokynů soustředil na přesně dané body v kokpitu, například ručičku výškoměru, nebo dané písmeno z imatrikulace letadla napsané na přístrojovém štítu. Díky zobrazování směru pohledu v reálném čase na monitoru je tak možné ihned vyhodnotit, zda je zařízení opravdu přesné již během měření. Pro potřeby této práce je navíc důležité rozpoznávání jednotlivých přístrojů, takže přesnost v řádu milimetrů není vyžadována. Z toho důvodu nebyly prováděny žádné pokusy a přesná určení odchylky měření, pouze vizuální ověření.

3.5 Metody zpracování dat

Výstupem z měření pokusu jsou data a videozáznamy, které musí být pro další použití exportovány s pomocí programu Pupil Player od společnosti PupilLabs. Před exportováním jsou v datových formátech, které není možné otevřít a čerpat z nich data. Po procesu zpracování pomocí tohoto softwaru je vytvořeno několik souborů a záznamů, které jsou již uživatelsky běžně přístupné. Po exportování jsou k dispozici 3 videozáznamy. Dva záznamy zobrazují pravé a levé oko, třetí obsahuje záznam scény. Program Pupil Player umožňuje širokou škálu modifikací a dodatečných informací, které lze přidat do těchto záznamů a následně je exportovat. Mezi tyto patří především vizualizace směru pohledu v reálné scéně, ale dále je například možné do rohu videa vložit i záznamy z kamer snímajících oči. Dále je možné modifikovat ukazatele směru pohledu, velikost bodu označujícího bod zájmu, trajektorii pohledu, nebo zobrazit oblast okolo bodu zájmu, která pokrývá plochu definovanou chybou měření v závislosti na použité metodě vyhodnocení eye trackingových dat. Pro následné zpracování pomocí softwaru na automatickou detekci objektů je využit čistý záznam, bez jakýchkoliv přidaných informací, dat a vizualizací. Konkrétně je využíváno pouze video z čelní kamery, snímající zorné pole uživatele. Postup zpracování dat je názorně ukázán na obrázku 3.5 a detailní popis všech kroků je popsán v následujících odstavcích.



Obrázek 3.5: Diagram postupu při zpracování dat.

Video záznam z čelní kamery, který byl pořízen, je v rozlišení 1280 x 720 px. Pro snížení časové náročnosti zpracování dat bylo nejprve nutné snížit rozlišení videa na polovinu. Pomocí funkce v programu Matlab tedy bylo rozlišení sníženo na hodnotu 640 x 360 px, a s tímto záznamem se dále pracovalo. Procesem zmenšení prošly záběry z čelní kamery ze všech měření. Ze tří záběrů, které dohromady tvořily zhruba 70 minut záznamu bylo náhodně vybráno 1500 snímků pro práci s detektorem objektů. Z každého záznamu bylo vybráno celkem 500 snímků. Na těchto snímcích bylo potřeba značit všechny oblasti zájmu, které má následně program detekovat. Označování probíhalo v prostředí programu Matlab 2021b, konkrétně s využitím aplikace Image Labeler z toolboxu Computer Vision. Vybrané oblasti zájmu byly letové, navigační a motorové přístroje případně větší oblast zájmu, obsahující více zařízení, které se týkají jedné oblasti jejich využití. Těmito oblastmi zájmu byly umělý horizont (AH), výškoměr (ALT), indikátor kurzu (HDG), indikátor vertikální rychlosti (VSI), GPS navigace (NAV), rychloměr (SPD), družný přístroj s koordinátorem zatačky a příčným sklonoměrem (TC), panel pro práci s rádiovým vybavení letadla (RTP), indikátor odchylky kurzu (CDI), automatický radiový kompas (ADF), plnicí tlak v sání motoru (MP), otáčky motoru (RPM) a stav hladiny paliva ve dvou palivových nádržích (Fuel). Jednotlivé oblasti byly označovány manuálně pro každý snímek, tento postup je znázorněn na obrázku 3.6.



Obrázek 3.6: Tvorba datasetu určeného pro trénování objektového detektoru.



Takto vytvořený dataset byl následně rozdělen na 1000 snímků pro fázi učení detektoru, a 500 snímků pro fázi ověření detektoru. Tyto snímky byly rozděleny tak, každý třetí snímek ze setu byl určen pro fázi ověření. Tím se zajistilo, že snímky byly vybrány opravdu náhodně ze všech videí. Detektor objektů byl vytvořen v prostředí programu Matlab 2021b. Byl navržen za využití metody R-CNN (Region Based Convolutional Neural Networks), konkrétně varianty Faster R-CNN. Tento přístup vyniká vyšší rychlostí detekce objektů, ale zachovává si dostatečnou přesnost a kvalitu. Fáze učení detektoru oblastí zájmu probíhala na architektuře Alexnet, která se skládá z pěti konvolučních a tří plně propojených vrstev. Pro samotné trénování byl využit algoritmus SGDM (Stochastic Gradient Descent with Momentum), který se pro trénování neuronových sítí často používá. Nastavení parametrů učení je znázorněno v tabulce 3.1 a hodnoty jsou určeny na základě předchozí zkušenosti s detektorem, který byl využíván pro podobné aplikace ve virtuální realitě.

Tabulka 3.1: Nastavení parametrů pro hluboké učení

Parametr	Hodnota
Momentum	0.9000
InitialLearnRate	1e-03
LearnRateSchedule	'piecewise'
LearnRateDropFactor	0.2000
LearnRateDropPeriod	1
L2Regularization	1.0000e-04
GradientTresholdMethod	'l2norm'
GradientTreshold	Inf
MaxEpochs	100
MiniBatchSize	2
Verbose	1
VerboseFrequency	50
Shuffle	'every-epoch'



SDGM je optimalizační algoritmus, který pro určení gradientu sestupu využívá určitý počet vzorků. Může se jednat buď o všechny vzorky v datasetu, což je efektivní pouze pro menší datasety, nebo naopak využívá jen jeden vzorek, což je velmi časově neefektivní. Metoda mini-batch gradient descent využívá k určení gradientu sestupu předdefinovaný počet vzorků, který je určen hodnotou *MiniBatchSize* (*mbs*), a v případě této práce nabýval hodnotu $mbs=2$. Proces učení byl vykonáván na GPU, konkrétně na NVIDIA GeForce RTX3070, protože GPU je schopna provádět tento typ výpočetních úloh efektivněji, než CPU.

Následně po ukončení procesu trénování proběhlo automatické testování detektoru objektů. K tomuto účelu sloužilo zbylých 500 označených snímků, které však nebyly využity pro trénování. Na základě porovnání výsledků detektoru a známých výsledků označených oblastí jsou pak vytvořeny PR křivky, sloužící k hodnocení detektoru. PR křivky (Precision/Recall) charakterizují závislost přesnosti a úplnosti predikce oblastí zájmu s pomocí detektoru. Tato závislost znamená, že čím více relevantních údajů má detektor rozpoznat, tím více falešně pozitivních detekuje zároveň s tím. Naopak, pokud má detektor detekovat pouze relevantní data bez falešně pozitivních, není schopen detekovat všechny relevantní, ale pouze jejich část. Tato závislost je popisována právě PR křivkami. Důležitým ukazatelem, který se využívá pro hodnocení přesnosti a správnosti je tzv. Intersection over Union (IoU). Tento ukazatel hodnotí, s jakou přesností dochází k překryvu detekované a předem známé oblasti. Matematicky se jedná o poměr průniku a sjednocení dvou oblastí zájmu. Pokud se oblast označená detektorem a oblast označená v testovacích datech 100% překrývají, je hodnota $IoU=1$. Standartně se využívá prahová hodnota $IoU=0,5$, při překročení hodnoty jde o úspěšnou detekci, při hodnotě nižší než prahové dochází k tzv. falešně pozitivní detekci. Při hodnocení může dojít k několika situacím spojeným s učením objektů.

- TP (True positive) je skutečně pozitivní určení oblasti zájmu. Nastává, když je hodnota $IoU \geq 0,5$



- FP (False positive) je falešně pozitivní určení oblasti zájmu. Nastává, když hodnota $IoU < 0,5$
- TN (True negative) je skutečně negativní určení oblasti zájmu
- FN (False negative) je falešně negativní určení oblasti zájmu

V případě TP dochází k tomu, že průnik detekované a testovací oblasti zabírá více než polovinu jejich sjednocené plochy. Detektor tedy správně detekuje danou oblast zájmu. Naopak v případě FP dochází k tomu, že detektor přiřadí určitou oblast k dané třídě, přesto že se na tom místě oblast náležící do dané třídy nenachází. Ukazatel TN se k hodnocení nevyužívá, ale jde o oblasti, kde není žádná oblast zájmu a detektor tam žádnou oblast nedetekoval. FN nastává ve chvíli, kdy detektor nedetekuje danou oblast zájmu, přestože by měl. Detektor je hodnocen s pomocí hodnot přesnosti (precision) a úplnosti (recall). Tyto ukazatele jsou definovány následujícími rovnicemi:

$$precision = \frac{TP}{TP + FP} \quad (3.1)$$

$$recall = \frac{TP}{TP + FN} \quad (3.2)$$

Po trénování a testování detektoru je další fází jeho aplikace na samotný záběr, který chceme analyzovat. Určování oblastí zájmu (ROI) proběhlo na všech třech záběrech a výsledkem detekce je matice, obsahující informace o detekovaných ROI pro každý snímek z videa. Tato matice obsahuje pro každý snímek záznamu názvy všech přístrojů, které byly v daný moment rozpoznány, současně se souřadnicemi rohových bodů pravidelných čtyřúhelníků definujících oblasti zájmu. Dalším možným výstupem z detektoru objektů je videozáznam obsahující veškeré detekované ROI. Vzhledem k tomu, že jednotlivé ROI jsou poměrně blízko sebe, dochází k překrývání okrajů oblastí a celý záznam se stává velmi nepřehledný. Snahou



této práce ale byla demonstrace zařízení a softwaru, proto byla snaha o co největší počet detekovaných oblastí zájmu.

Další data, která vyžadují určité zpracování, jsou údaje o pohybech očí získané při měření pokusu. Ty můžeme obecně rozdělit do dvou kategorií. Primárně jde o data, týkající se polohy zornic a oka, a druhý typ dat nám popisuje polohu výsledného bodu pohledu ve vztahu k záznamu z čelní kamery. Data o směru pohledu nejsou určována přímo měřením, ale jsou dopočítávána z polohy zornic na základě kalibrace zařízení. Pro potřeby práce s detektorem objektů se využívají data o směru pohledu. Pracuje se s normalizovanými hodnotami souřadnic x a y . Teoretický rozptyl hodnot je neomezený v každém směru, ale obraz z čelní kamery zabírá pouze část zorného pole a je popsán souřadným systémem se středem v pravém horním rohu s rozsahem hodnot 0-1 na obou osách. Pokud je pohled směřován do zorného pole zabíraného čelní kamerou, hodnoty se pohybují ve zmiňovaném rozmezí, ale pokud je pohled pod větším úhlem směřován mimo dosah kamery, mohou se v datech objevit i hodnoty záporné, nebo vyšší než 1.

Vzhledem k tomu, že cílem je videozáznam čelní kamery a údaje o směru pohledu synchronizovat, a pohled do scény vizualizovat, musí být souřadnice směru pohledu škálovány na stejné rozlišení, jako má záznam. Rozlišení videa se ale v prvním kroku zmenšilo, proto je potřeba i eye trackingová data škálovat na rozlišení 640 x 360 px. To se provede prostým roznásobením normalizovaných hodnot souřadnic s hodnotou rozlišení videa. Poté následuje úprava dat tak, aby souhlasily časové známky videa a ET. Video má obnovovací frekvenci 30 Hz, zatímco údaje o poloze se sbírají s frekvencí zhruba 200 Hz. Bylo proto nutné vybrat pro každý frame videa pouze jeden odpovídající záznam o poloze očí, čímž zároveň došlo k zásadnímu zmenšení celkového objemu dat, se kterým se pracuje. Takto vyselektovaná data se dají promítnout přes videozáznam scény a tím se zobrazuje bod, kam je v daný moment směřován pohled. Záměrem ale bylo opatřit tento bod popiskem, který by navíc identifikoval danou oblast zájmu, a zároveň by onen přístroj zviditelnil rozsvícením jeho detekovaných hranic.



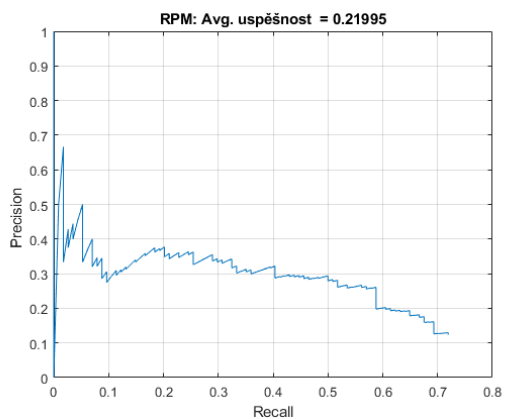
Přiřazení směru pohledu k ROI probíhalo pomocí indikace, jestli se souřadnice bodu zájmu nachází v nějaké oblasti zájmu. Tyto oblasti jsou z detektoru definovány jako pravidelné čtyřúhelníky a jsou popsány s pomocí souřadnic jejich rohů. Na základě těchto souřadnic byly vytvořeny polygony detekovaných oblastí zájmu a bylo vyhodnocováno, zda se souřadnice pohledu nachází uvnitř těchto polygonů. V případě pozitivního výsledku pak byl k bodu zájmu přiřazen štítek s odpovídajícím označením ROI.



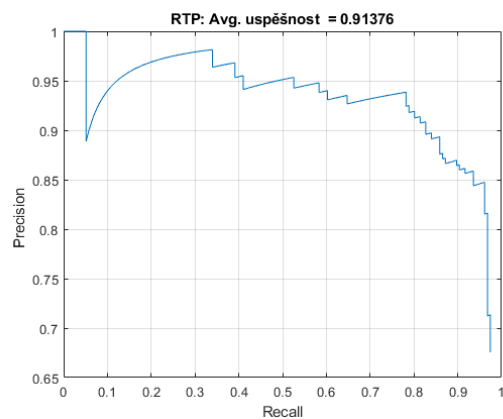
4 Prezentace výsledků

V průběhu zpracování praktické části této práce vzniklo několik dílčích výsledků, které se dají použít již ve své aktuální podobě, nebo se mohou zpracovávat dále. Kombinací těchto dílčích výsledků se pak dá dosáhnout různých finálních výstupů. Prvotní fází zpracování proběhla fáze učení detektoru objektů v kokpitu. Výkonnost detektoru se hodnotí pomocí PR křivek, které udávají průměrnou úspěšnost detekce a závislost přesnosti a úplnosti detekovaných dat. Na obrázku 4.1, 4.2 a 4.3 jsou znázorněny PR křivky pro testování detekce všech oblastí zájmu. Hodnota úspěšnosti detekce, dále označovaná proměnnou p , se u jednotlivých přístrojů lišila v závislosti na mnoha faktorech. Nízkou průměrnou úspěšnost vykazují obecně přístroje, které se nachází v okrajové části zorného pole čelní kamery, nebo jsou ve větší vzdálenosti od pozice pilota. Nejnižší hodnoty vykazoval indikátor otáček motoru, viz obrázek 4.1a. Tento přístroj byl ve většině případů na okraji zorného pole, byl ve větší vzdálenosti než ostatní přístroje, navíc ho v mnoha případech zakrývaly berany řízení. Vzorek označených dat pro učení detektoru proto nebyl dostatečný a výsledná průměrná úspěšnost určení dosahovala pouze hodnoty $p=0,22$. Pro srovnání, přístroj GPS navigace se nachází blíž pilotovi, je dobře viditelný na většině snímků, i v případě zhoršených světelných podmínek. Navíc tento displej zabírá mnohem větší oblast než klasické budíky. V případě navigace se proto hodnota úspěšnosti pohybovala okolo $p=0,89$. Ještě vyšší úspěšnosti určení dosáhl panel s radiovými přístroji, konkrétně $p=0,91$. PR křivka tohoto přístroje je zobrazena na obrázku 4.1b. Tyto vyšší hodnoty jsou dány především větším rozměrem přístroje, tím pádem větší šancí pro detektor na správné určení.

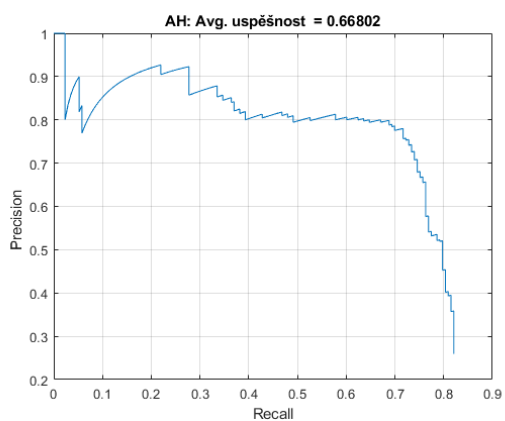
U přístrojů, které se nachází přímo před pilotem a nejsou zakrývány jinými objekty, se dosahovalo průměrné úspěšnosti v rozmezí 0,6 – 0,7. Drobnou odchylkou od těchto hodnot byl rychloměr, který byl určován s úspěšností $p=0,54$. Detekce umělého horizontu probíhala s úspěšností $p=0,67$, u výškoměru byla hodnota $p=0,64$, u ukazatele vertikální rychlosti $p=0,64$ a u indikátoru kurzu byla hodnota $p=0,67$. Rychloměr byl určován s úspěšností $p=0,54$, koordinátor zatáčky s hodnotou $p=0,67$, indikátor odchylky kurzu měl úspěšnost $p=0,61$



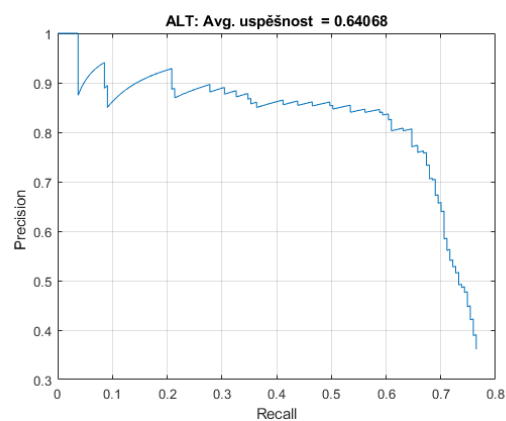
(a)



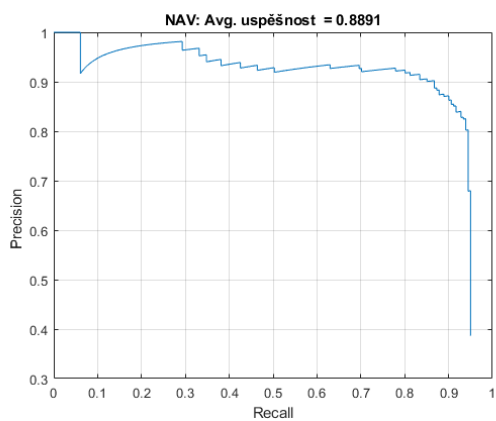
(b)



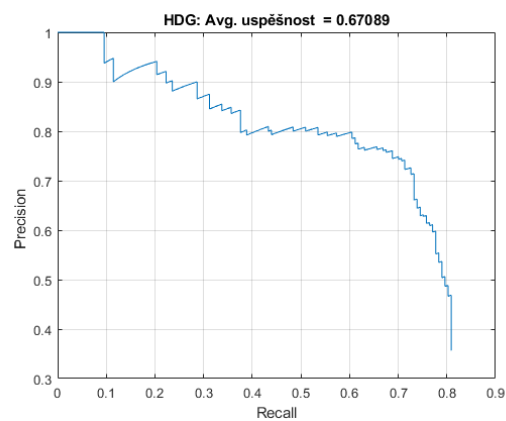
(c)



(d)

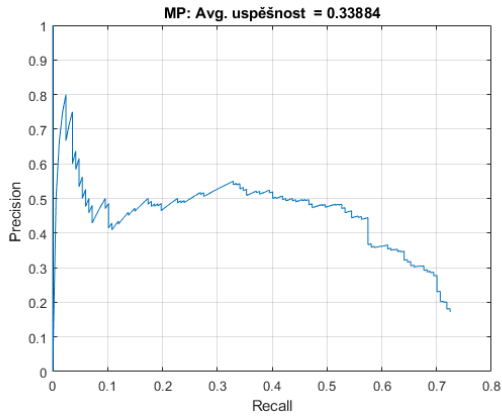


(e)

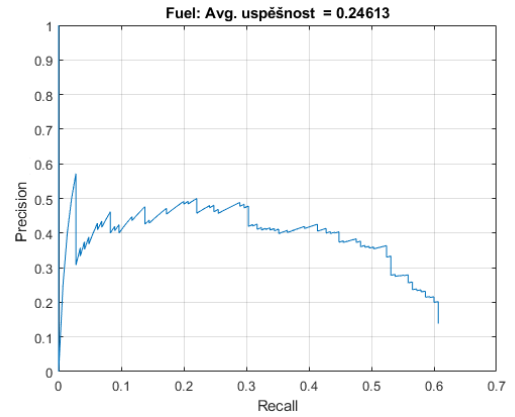


(f)

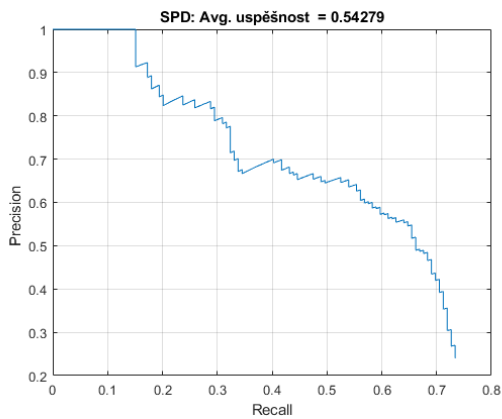
Obrázek 4.1: PR grafy úspěšnosti detekce na testovacím datasetu pro objekty RPM (a), RTP (b), AH (c), ALT (d), NAV (e) a HDG (f).



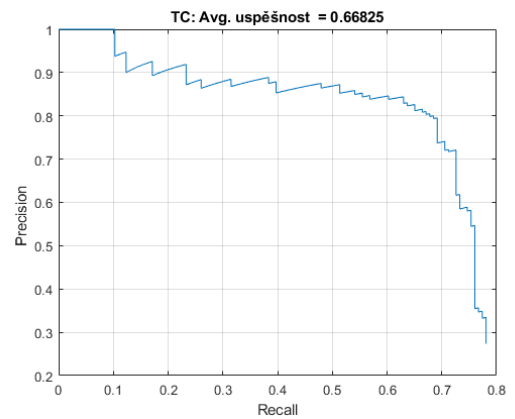
(a)



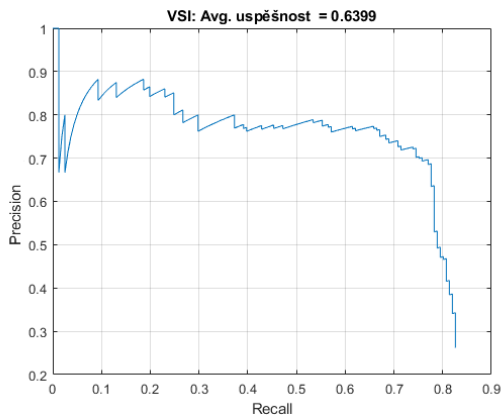
(b)



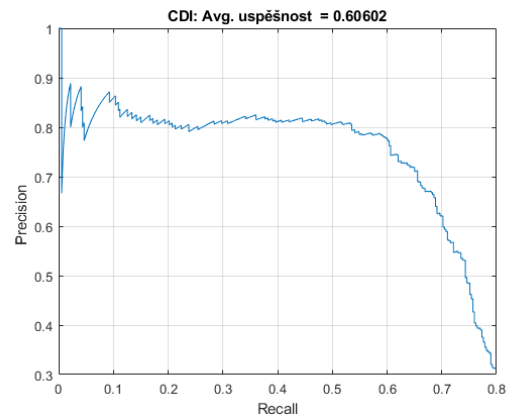
(c)



(d)

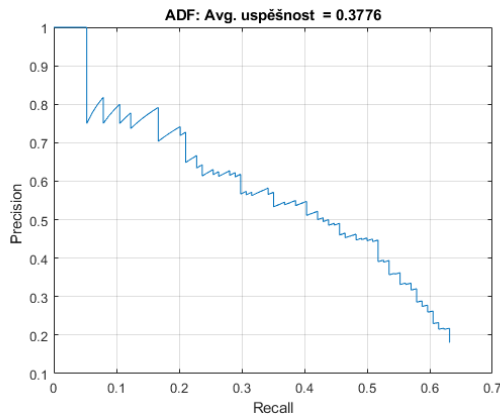


(e)



(f)

Obrázek 4.2: PR grafy úspěšnosti detekce na testovacím datasetu pro objekty MP (a) a Fuel (b) SPD (c), TC (d), VSI (e) a CDI (f).



(a)

Obrázek 4.3: PR graf úspěšnosti detekce na testovacím datasetu pro objekt ADF (a).

a u automatického radiového kompasu byla $p=0,38$. Nejvyšší úspěšnost určení vykazovaly přístroje NAV a RTP na středovém panelu, kvůli svým rozměrům a stále ještě výhodnému umístění v zorném poli. U přístroje navigace byla úspěšnost detekce $p=0,89$ a u radiového panelu $p=0,91$. Naopak nejnižší úspěšnost detekce nastala u přístrojů MP, RPM a Fuel, u kterých se kombinuje několik nepříznivých faktorů. Vyskytují se daleko pro detekci pomocí kamery, jsou částečně zakryté jinými objekty, a jedná se o poměrně malé přístroje, což by jejich detekci zhoršovalo i v ideálních podmínkách. U přístroje idnikujícího plnící tlak byla úspěšnost $p=0,34$, u otáčkoměru byla $p=0,22$ a u ukazatele stavu paliva byla úspěšnost $p=0,25$. Vizualizace všech detekovaných ROI je znázorněna na obrázku 4.4. Program vykreslí veškeré oblasti, které se mu podařilo rozpoznat, včetně údaje o důvěryhodnosti určení. Každá jednotlivá zóna tak má vlastní údaj *confidence*, a je opatřena štítkem s třídou, do které oblast zájmu spadá.

Na obrázku 4.5 je znázorněna finální podoba detektoru tak, jak byl navržen. Pro větší přehlednost bylo zvoleno zobrazení pouze takového ROI, na který je v daný moment soustředěn pohled. Vizualizace tedy neobsahuje všechny detekované oblasti zájmu, jak bylo dříve uvedeno, ale pro větší přehlednost pouze tu aktuální. Směr pohledu je vizualizován s pomocí červeného bodu v obrazu, který je navíc opatřen popiskem odpovídajícím třídě oblasti

zájmu. Každý ROI je navíc při detekci zvýrazněn pomocí barevného světlého čtverce. Díky tomu, že je v procesu vizualizace výstupu přiřazen pohled v každém snímku videa k danému ROI, je možné časově synchronizovaná data exportovat a získat kvalitativní i kvantitativní data o pohledu pilota.

Pro porovnání výsledků je na obrázku 4.6 znázorněn i případ, kdy za letu nastaly nepříznivé světelné podmínky a nebyla dobrá viditelnost kamery na přístrojový štít. V takové případě je vidět, že systém nemůže detekovat žádné, nebo jen omezené množství oblastí zájmu



Obrázek 4.4: Vizualizace všech detekovaných oblastí zájmu v daném snímku.



Obrázek 4.5: Výsledná podoba detektoru objektů, obsahující indikaci ROI, kam aktuálně směřuje pozornost.



Obrázek 4.6: Příklad špatných světelných podmínek, kdy detekce objektů je neuspokojivá.



5 Diskuze výsledků

Pro sestrojení nositelného řešení byl v této práci využit již existující hardware, proto se od samého začátku nepočítalo s tím, že by právě hardware mohl být jednou z největších limitací práce. Průběh experimentu nám ukázal, že zařízení funguje velmi uspokojivě za příznivých světelných podmínek a při správném zacházení a používání. Jakmile však došlo ke zhoršení světelných podmínek, systém začal vykazovat hned několik nedokonalostí, vázaných ke všem kamerám, kterými jsou brýle osazeny.

Jednou z limitací, která se konkrétně týká čelní kamery, je její nízký dynamický rozsah světelných podmínek. Jedná se o základní senzor, ne nijak kvalitní, o tom svědčí na dnešní dobu nízké rozlišení 1280x720 pixelů. Jakmile je snímaná scéna rozdělena na více částí, které se liší svou barvou, nasvícením, nebo je mezi nimi ostrý světelný přechod, kamera se dokáže přizpůsobit pouze na jedny světelné podmínky. Vlivem toho je většinu času velmi světlý výhled ven z letadla, zatímco pohled na přístrojový štít je velmi tmavý, bez možnosti rozeznat jednotlivé přístroje. Pro přizpůsobení se tmavosti štítu je tak potřeba, aby tato tmavá zóna převažovala v obrazu scény a kamera se přizpůsobila právě na tmavší světelné podmínky. V tuto chvíli však zase ztrácíme výhled ven, který je příliš světlý.

Druhá limitace se také týká světelných podmínek a kamer pro snímání očí. Tyto kamery pracují v infračerveném spektru elektromagnetického vlnění. Pokud je obličej přímo osvětlen silným světlem, nedokážou tyto kamery upravit svoje parametry a kompenzovat světlejší scénu. V důsledku toho je zhoršená viditelnost kamer a následná detekce zornic. Tyto kamery navíc disponují rozlišením pouze 192x192 pixelů, které představuje další limitaci. Při intenzivním světle zornice reagují svým zmenšením, které je dělá pro kamery ještě hůře detekovatelné. Díky tomu, že všechny tyto limitace souvisí s intenzivním světlem, tak se za ideálních podmínek nevyskytují téměř vůbec, nebo naopak při velmi nepříznivých podmínkách se vyskytují všechny najednou. To velmi ztěžuje používání zařízení právě v popsaných podmínkách intenzivního světla. Kamera snímající scénu nám po většinu



času nedokáže poskytnout relevantní záznam kokpitu a data o pohybech očí jsou často nespolehlivá, hodnota konfidence se pohybuje pod úrovní 0,6.

Limitací navrženého řešení je také fakt, že je založeno na obroučkách slunečních brýlí bez slunečních skel. To fakticky neumožňuje nošení dioptrických ani jiných slunečních brýlí. Především v případě dioptrických brýlí se může jednat o limitující faktor. Řešením mohou být kontaktní čočky, které funkci systému nijak neovlivňují a mohou být bez problému používány. Adekvátní ochrana proti silnému slunci pak není s takto navrženým řešením možná, nicméně se nejedná o problém, který by přímo znemožňoval bezpečné používání brýlí za letu.

Jednou z limitací systému, kterému se věnuje tato práce, je samotný hardware pro sledování pohybu očí. Zařízení musí být spojeno kabelem s počítačem, na kterém se provádí záznamy ze všech kamer a jejich následné zpracování. Pokud by mělo být zařízení bezdrátové, byla by jeho použitelnost v řádu minut. Proto prototyp navržený v této práci vyžaduje připojení k PC. Délka kabelu však může být libovolná, omezení v používání je proto pouze velmi malé. Navíc zamýšlený způsob využití v kokpitu umožňuje vedení přípojného kabelu komfortně tak, aby neomezoval při vykonávání činností na palubě.

Během hodnocení průměrné přesnosti detektoru jsem se v práci potýkal s nižšími hodnotami, než bylo původně očekáváno. Hodnoty se pohybovaly často nad hranicí úspěšnosti 60 %. Částečně by se dalo lepších výsledků dosáhnout objemnějším datasetem označených oblastí zájmu. Další možností by bylo využití videa ve vyšším rozlišení, pro větší míru detailu přístrojů. V tomto případě je však limitující i výpočetní výkon počítače, který data zpracovává, a vyšší rozlišení by násobně prodloužilo dobu zpracování. Tyto možnosti by však pravděpodobně samostatně nevedly k zásadnímu zlepšení výsledků. Detekované oblasti zájmu totiž svou plochou v celkové scéně často zabíraly velmi malou část. Když pak detektor určil oblast zájmu a došlo jen k částečnému překryvu známé a detekované oblasti zájmu, byl daný vzorek vyhodnocen jako negativní. U přístrojů, které byly větší, jako například GPS navigace byla úspěšnost určení přes 90 %. Zároveň bylo horších výsledků přesnosti dosaženo u přístrojů, které se nacházely většinu času na okraji zorného pole, případně byly částečně



skryty za berany řízení. Trénovací data pro tyto přístroje tak nebyla dostatečná a kvůli tomu i výsledná detekce byla na nižší úrovni než u přístrojů v přímém zorném poli.



6 Závěr

Cílem této práce bylo vytvoření nositelného řešení pro detekci pohybu očí a zároveň vytvoření softwarového řešení pro zpracování dat získaných pomocí navrženého zařízení. Pro dosažení takových cílů bylo vytvořeno zařízení, které je založeno na dostupném materiálu a využívá kamery, které jsou k dispozici na ústavu letecké dopravy fakulty dopravní při ČVUT. Nosičem kamer jsou obroučky slunečních brýlí a jejich uchycení je vyřešeno pomocí platových dílů vytvořených s pomocí 3D tisku. Experimentální měření probíhalo nejen za ideálních podmínek, ale i v reálném prostředí během letu. Při měření v ideálních podmínkách bylo dosaženo kvalitních výsledků s vysokou pravděpodobností správnosti měření. Během provádění pokusu za letu se výsledky bylo dosaženo uspokojivých výsledků, ale pravděpodobnost správnosti měření silně závisela na okolních světelných podmínkách. Na základě těchto poznatků pak byly určeny limitace zařízení. Při zohlednění cenové kategorie kamer, které byly použity, se však jedná o předpokládané výkony.

Během přípravy prototypu se vyskytlo několik problémů, především v oblasti ergonomie zařízení. Kvůli tomu muselo být vytištěno více modelů a postupným vylepšováním jejich tvaru a geometrie uchycení kamer se dosáhlo finální podoby, která vyhovuje požadavkům na komfortní nošení během delšího časového úseku.

Součástí práce bylo vytvoření datasetu s označenými oblastmi zájmu pro kokpit letounu Cessna 182, kde toto zařízení bude fungovat bez dalšího trénování detektoru. Pro využití v jiném typu kokpitu by bylo potřeba vytvořit set označených dat, postup pro tento případ je však v práci podrobně popsán. Výstup detektoru poskytuje vizualizaci eye trackingových dat, včetně přiřazení odpovídající oblasti zájmu. Tato data je navíc možné exportovat i do datové podoby, kde ke každému snímku je přiřazena oblast zájmu, pokud byla nějaká detekovaná a byla na ní soustředěna pozornost.

Vzhledem k tomu, že sledování pohybu očí je velmi podstatnou součástí velkého množství prací, které se zabývají interakcí člověka s okolím, je zde předpoklad, že by se nositelné



řešení navržené v této práci mohlo v budoucnu využít. Ceny srovnatelných produktů na trhu jsou řádově vyšší a ceny kvalitnějších výrobků se pohybují v řádech statisíců. Toto zařízení může sloužit jako ideální prostředek pro získávání cenných dat v budoucích pracích, které se navíc nemusí týkat letectví. Využití zařízení není limitováno oborem. Díky doplnění systému o software na zpracování dat a detekci objektů navíc může sloužit k vyhodnocování a klasifikaci naměřených dat. Tato práce ve výsledku prokázala, že lze využít zařízení s původně jiným určením a transformovat ho pro využití v reálném prostředí při zachování všech původních vlastností produktu.



Seznam použité literatury

- [1] George K. Mikros and Ján Macutěk, editors. *Sequences in Language and Text*. De Gruyter Mouton, 2015.
- [2] Hazel I. Blythe, Simon P. Liversedge, Holly S.S.L. Joseph, Sarah J. White, John M. Findlay, and Keith Rayner. The binocular coordination of eye movements during reading in children and adults. *Vision Research*, 46(22):3898–3908, 2006.
- [3] Bart Krekelberg. Microsaccades. *Current Biology*, 21(11), 2011.
- [4] Casper J. Erkelens. Coordination of smooth pursuit and saccades. *Vision Research*, 46(1-2):163–170, 2006.
- [5] Jaromír Hanuška, Cecilia Bonnet, Jan Ruzs, Tomáš Sieger, Robert Jech, Sophie Rivaud-Péchoux, Marie Vidailhet, Bertrand Gaymard, and Evžen Růžička. Fast vergence eye movements are disrupted in parkinson’s disease. *Parkinsonism & Related Disorders*, 21(7):797–799, 2015.
- [6] V. P. Laurutis and D. A. Robinson. The vestibulo-ocular reflex during human saccadic eye movements. *The Journal of Physiology*, 373(1):209–233, 1986-04-01.
- [7] E. M. Ornitz, M. B. Brown, Anne Mason, and N. H. Putnam. The effect of visual input on post-rotatory nystagmus in normal children. *Acta Oto-Laryngologica*, 77(1-6):418–425, 2009-07-08.
- [8] Mark Shelhamer and Dale C. Roberts. Magnetic scleral search coil. In *Vertigo and Imbalance: Clinical Neurophysiology of the Vestibular System*, pages 80–87. Elsevier, 2010.
- [9] Muhammad Syaiful Amri bin Suhaimi, Kojiro Matsushita, Minoru Sasaki, and Waweru Njeri. 24-gaze-point calibration method for improving the precision of ac-eog gaze estimation. *Sensors*, 19(17), 2019.



- [10] Joseph M. Furman and Floris L. Wuyts. Vestibular laboratory testing. In *Aminoff's Electrodiagnosis in Clinical Neurology*, pages 699–723. Elsevier, 2012.
- [11] Konrad P. Körding, Christoph Kayser, Belinda Y. Betsch, and Peter König. Non-contact eye-tracking on cats. *Journal of Neuroscience Methods*, 110(1-2):103–111, 2001.
- [12] Randall L. Harris and David M. Christhilf. What do pilots see in displays? *Proceedings of the Human Factors Society Annual Meeting*, 24(1):22–26, 1980.
- [13] Erland A. I. Svensson and Glenn F. Wilson. Psychological and psychophysiological models of pilot performance for systems development and mission evaluation. *The International Journal of Aviation Psychology*, 12(1):95–110, 2009-11-13.
- [14] Zbigniew Gomolka, Boguslaw Twarog, Ewa Zeslawska, and Damian Kordos. Registration and analysis of a pilot's attention using a mobile eyetracking system. In *Engineering in Dependability of Computer Systems and Networks*, pages 215–224. Springer International Publishing, Cham, 2020.
- [15] Nadine B. Sarter, Randall J. Mumaw, and Christopher D. Wickens. Pilots' monitoring strategies and performance on automated flight decks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(3):347–357, 2007.
- [16] Kishan Mehrotra, Chilukuri K. Mohan, and Sanjay Ranka. *Elements of Artificial Neural Networks*. MIT Press, Cambridge, MA, USA, 1996.
- [17] Marc Assens, Kevin McGuinness, Xavier Giró-i Nieto, and Noel O'Connor. Scan-path prediction on 360 degree images using saliency volumes. 07 2017.
- [18] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <<http://www.deeplearningbook.org>>.
- [19] Hamed Habibi Aghdam and Elnaz Jahani Heravi. Traffic sign detection and recognition. In *Guide to Convolutional Neural Networks*, pages 1–14. Springer International Publishing, Cham, 2017.



- [20] Muhamad Yani, M.T. Budhi Irawan S, Si., and M.T. Casi Setiningsih S.T. Application of transfer learning using convolutional neural network method for early detection of terry's nail. *Journal of Physics: Conference Series*, 1201(1), 2019-05-01.
- [21] Geoffrey Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint*, arXiv, 07 2012.
- [22] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778. IEEE, 2016.
- [23] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015.
- [24] Guoxu Liu, Joseph Christian Nouaze, Philippe Lyonel Touko Mbouembe, and Jae Ho Kim. Yolo-tomato. *Sensors*, 20(7), 2020.