

Rozpoznávanie biometrických údajov pomocou neurónových sietí – identifikačné znaky ucha

Gabriela Andrejková, Peter Kažimír

Ústav informatiky, Prírodovedecká fakulta UPJŠ
Jesenná 5, Košice

E-mail: Gabriela.Andrejko@upjs.sk, Peter.Kazimir@gmail.com

Abstrakt

Ucho človeka, jeho veľkosť, tvar je pravdepodobne jedinečné pre každého človeka. Toto tvrdenie nebolo ani dokázané ani vyvrátené. Bolo urobených niekoľko experimentov, ktoré identifikujú ľudí podľa uší, ak sú obrázky ich uší už v databáze. Tieto systémy pracujú veľmi spoľahlivo a vyššia ako 90 percentná účinnosť nie je ojedinelá. V príspevku je popísaný nový systém používajúci konvolučné neurónové siete na identifikáciu štyroch charakteristických bodov ucha. Konvolúcia tu súvisí so zameraním sa na detaily. Systém bol odskúšaný na existujúcich databázach obrázkov uší, ale tiež boli vytvorené ďalšie vlastné obrázky. Dosiahnuté výsledky pri experimente - klasifikácia na 34 obrázkoch uší - ukazujú, že systém je použiteľný na klasifikáciu uší podľa skúmaných znakov, podľa ktorých vznikne malá trieda uší, ktorú už nie je problém systematicky prehľadať.¹

1 Úvod

Niektorí kriminalisti sa domnievali, že ľudské ucho je jedinečné pre každého človeka a nemenné po dovŕšení dospelosti. Niektorí z nich tvrdia, že charakteristické znaky ucha sa dajú použiť pri identifikácii osoby, ako sa napríklad používajú odtlačky prstov. U niektorých pretrváva názor, že ponechaný odtlačok ucha na mieste činu, sa dá identifikovať porovnaním s odtlačkom ucha podozrivého. Tieto názory sa zdajú byť pravdivé, ale neboli nikdy dokázané. V súčasnosti pravdepodobne neexistuje ani žiadna vedecká štúdia o tom, že ľudské uši sú zreteľne odlišné. Možno hlavnou príčinou toho je, že ľudia, ktorí toto skúmali, nemali dostatok skúsenosti na vytvorenie štúdie, ktorá by to dokázala. Väčšina ľudí, ktorá sa tým zaoberala, nepracovala vo vedeckej sfére a svoje skúsenosti nadobudli pri práci na rôznych prípadoch. Napriek tomu existuje viacero príspevkov týkajúcich sa identifikácie človeka pomocou znakov ucha a je to celkom zaujímavá problematika.

Ďalšie členenie príspevku bude nasledujúce: V druhej kapitole popíšeme prehľad metód, ktoré boli používané

v tejto oblasti. V tretej kapitole sa budeme venovať stručnému popisu konvolučných neurónových sietí (KNS) a popisu ich vlastností. Štvrtá kapitola bude venovaná popisu novovytvoreného systému na detekciu základných bodov ucha, v piatej kapitole popíšeme experimentálne výsledky dosiahnuté použitím navrhnutého systému. V závere uvedieme odporúčania pre použitie navrhnutého systému.

2 Prehľad metód a výsledkov

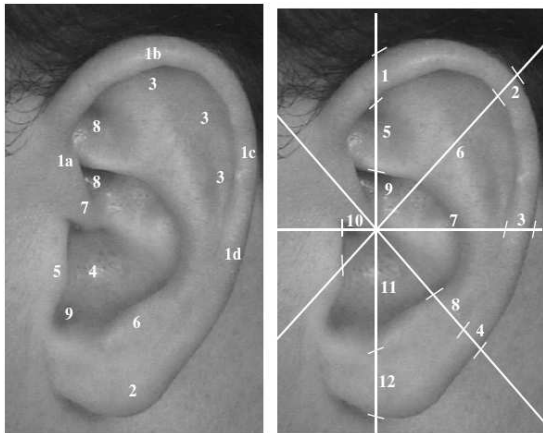
R. Imhofer (1906) [2] – pražský doktor, urobil dôležitý krok k tomu, aby sa ucho stalo významným faktorom pri identifikovaní človeka. Prišiel na to, že vo vzorke 500 uší potreboval len 4 charakteristiky na ich jednoznačné rozlíšenie.

Ch. Fields (1960) [1] – urobil pokus, pri ktorom sa snažil identifikovať novonarodené bábätká na základe uší. Študoval vzorku 206 uší a usúdil, že ucho novorodenca má dostatok parametrov k tomu, aby vedel odlíšiť jedno bábätko od druhého.

A. Iannarelli (1989) [5] - vyvinul antropometrickú techniku identifikácie založenú na biometrii ucha. "Iannarelliho systém" je založený na 12 meraniach ilustrovaných na obrázku 1. Zobrazené miesta sú merané na špeciálne zarovnaných a normalizovaných fotografiách ľavého ucha. Normalizovať a zarovnať obrázky znamená, že sú projektované na štandardný "Iannarelliho" zväčšený podstavec, ktorý je posúvaný horizontálne a vertikálne, až kým sa obrázok ucha dostane do predpísanej polohy. Tento systém si vyžaduje presné zarovnanie a normalizáciu fotografií ucha.

Iannarelli použil systém aj pri dvoch svojich veľkých experimentoch. V prvom experimente porovnával 10.000 uší náhodne vytiahnutých z databázy v Kalifornii a v druhom experimente porovnával identické dvojčatá a trojčatá, (ktoré majú rovnaký genetický základ). V oboch experimentoch došiel k záveru, že uši sú jednoznačné a dokonca aj dvojčatá a trojčatá sa dajú podľa toho rozlíšiť.

¹ Výskum je podporovaný VEGA grantom č. 1/0035/09



Obr. č. 1. Anatómia ucha a merania na uchu

B. Victor (2002) [6] – urobil experiment, pri ktorom porovnával výkonnosť metódy Principal Component Analysis (PCA) pri identifikácii tváre a ucha. V experimente použil 294 subjektov, u ktorých bral do úvahy aj faktory ako vek, osvetlenie, výraz tváre a čas, ktorý uplynul medzi dvoma fotografiami. Jeho výsledky ukázali, že PCA je výkonnejšia pri identifikácii tváre ako pri identifikácii ucha.

K. Chang (2003) [7] – urobil rovnaký experiment ako Victor, ale s rôznym výsledkom. Jeho experiment dokázal, že medzi použitím PCA metódy na ucho a na tvár nie je až taký rozdiel. Rozpoznávací pomer pri identifikácii ucha dosiahol 71.6 percenta a pri tvári dosiahol tento ukazovateľ 70.5 percenta. Chang sa domnieva, že dôvodom prečo Victor došiel k takému záveru je, že v obrázkoch, ktoré mal k dispozícii si nedal pozor na prípadné rušivé momenty akými sú napr. náušnice, vlasy, svetlo atď. Chang urobil aj experiment, pri ktorom použil pri identifikácii obidva biometrické údaje (tvár a ucho) tzv. multimodálne porovnávanie. Pri tomto experimente dosiahol rozpoznávací pomer 90.9 percenta

P. Yan, K. W. Bowyer [8] – v roku 2005 urobili najväčší experiment automatického rozpoznávania ucha. Najväčší v zmysle počtu osôb, obrázkov a použitých algoritmov. V experimente boli použité 3D obrázky od 404 osôb, na ktoré sa aplikovali štyri algoritmy aplikujúce rôzne metódy. Pri tomto experimente skúmali 3 metódy na vybratie orientačných bodov ucha.

Existuje mnoho ďalších príspevkov zaoberajúcich sa touto problematikou [12, 13, 14, 15, 16, 17]. V novších príspevkoch boli pri riešení použité tiež neurónové siete (NS) a genetické algoritmy (GA). V tabuľke 1 uvedieme

stručný prehľad použitých metód a dosiahnutých výsledkov.

Tab. 1. Zhrnutie niektorých metód a výsledkov pri identifikácii uši.

Konatelia exper.	Počet vzoriek, Dáta	Aplik. metóda	Náušnice /Čiastočné Zahalenia	Dosiahnutý pomer
Chen & Bhanu [14]	52, 3D	ICP	NIE	90.4%
Hurley [15]	252, 2D	PCA	NIE	99.2%
Moreno [16]	28, 2D	NS	Nespom.	93%
Yuizono [17]	110, 2D	GA	Nespom.	100%
Victor & Chang [6.7]	88, 2D	PCA	NIE	73%
Yan & Bowyer [8]	404, 3D	ICP	NIE	97.5%

3 Konvolučné neurónové siete

Problémy, s ktorými sa môžeme stretnúť u vrstvových sietí, sú riešené u konvolučných neurónových sietí (KNS). KNS sa automaticky naučia extrahovať lokálne znaky, sú stále a nemenné voči malým skresleniam a lokálnym deformáciám vo vstupnom vzore a implementujú princíp zdieľania váh, čo vo významnej miere redukuje počet nastaviteľných parametrov a zvyšuje ich použiteľnosť. LeCun [9] prezentoval prvú konvolučnú neurónovú sieť, ktorá bola tréňovaná učiacim algoritmom Backpropagation a bola použitá pri rozpoznávaní rukou písaných číslíc.

3.1 LeCunov model konvolučnej siete

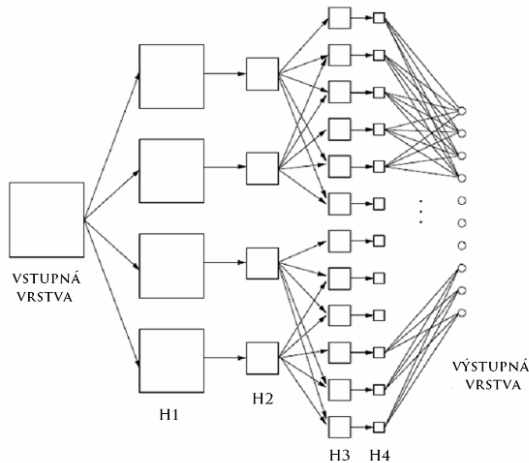
Na obrázku č. 2 je zobrazená KNS, nazvanú *LeNet-1*, ktorú navrhol a použil LeCun [9]. Popis jej štruktúry:

- Na vstupnú vrstvu $y^{(0)}$ je prezentovaný čierno-biely obrázok, ktorý má rozmery 28 x 28 pixlov. Intenzita farby pixlov je normalizovaná medzi -1 a +1.
- Prvá skrytá vrstva H1 pozostáva zo 4 znakových máp $y_j^{(1)}$ a odchýlky $b_j^{(1)}$, $j=1, 2, 3, 4$. Každá zo 4 znakových máp má 25 váh $w_{j0}^{(1)}(u,v)$, ktoré tvoria 5 x 5 tréňované jadro. Hodnoty znakovkej mapy $y_j^{(1)}(x,y)$ sa vypočítajú konvulciou vstupnej mapy $y^{(0)}$ s príslušným jadrom $w_{j0}^{(1)}$ a aplikovaním aktivačnej funkcie $\theta^{(1)}$ na výsledok:

$$y_j^{(1)}(x,y) = \theta^{(1)} \left(\sum_{(u,v) \in K} w_{j0}^{(1)}(u,v) y^{(0)}(x+u, y+v) + b_j^{(1)} \right)$$

$$\text{kde } K = \{ (u,v) \in N^2 \mid 0 \leq u < 5, 0 \leq v < 5 \}.$$

Pri tomto treba poznamenať, že kvôli efektu okrajov budú znakové mapy menšie, t. j. 24 x 24. Každá konvolučná mapa je nasledovaná priemerovacou mapou $y_j^{(2)}$, ktorá vykonáva určitý typ priemerovania a redukovania dimenzie príslušnej konvulčnej mapy o polovicu:



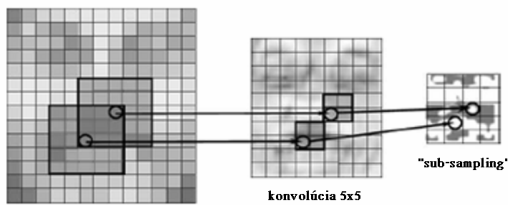
Obr. č. 2 Konvulčná neurónová sieť

- Priemerovacie mapy vrstvy H2 majú rozmer 12 x 12. Každá priemerovacia mapa má svoju váhu $w_j^{(2)}$ a odchýlku $b_j^{(2)}$. Teda pre hodnoty priemerovacej mapy bude platiť:

$$y_j^{(2)}(x, y) = \theta^{(2)} \left(w_j^{(2)} \times \sum_{(u,v) \in \{0,1\}^2} y_j^{(1)}(2x+u, 2y+v) + b_j^{(2)} \right)$$

kde $\theta^{(2)}$ je opäť aktivačná funkcia. Obrázok č. 3 znázorňuje proces konvulcie a tvorby menších vzoriek.

- Vrstvy H3 a H4 obsahujú 12 konvulčných a 12 priemerovacích máp. Konvulčné majú rozmer 8 x 8 a priemerovacie 4 x 4 pixlov.



Obr. č. 3 Vstupný obrázok nasledovaný znakovou mapou, vyrobenou konvulciou s 5 x 5 konvulčným jadrom, a 2 x 2 priemerovacou mapou.

Funkcia, ktorú využívajú vrstvy H3 a H4, je rovnaká ako pre vrstvy H1 a H2, s tým rozdielom, že pre vrstvu H3 sa konvulcia vykonáva s 3 x 3 konvulčným jadrom namiesto 5 x 5. Ďalej konvulčná mapa j vo vrstve H3 môže mať niekoľko jadier $w_{ji}^{(3)}$ operujúcich na rôznych mapách i z predchádzajúcej mapy H2. Aktivácia príslušnej konvulčnej mapy je získaná sčítaním výsledkov z konvulcií a odchýliek. Vo všeobecnosti aktivačná formula pre konvulčnú mapu j vo vrstve l je:

$$y_j^l(x, y) = \theta^{(l)} \left(\sum_{i \in I} \sum_{(u,v) \in K} w_{ji}^l(u, v) y_i^{(l-1)}(x+u, y+v) + b_j^{(l)} \right)$$

$$\text{kde } K = \{ (u, v) \in \mathbb{N}^2 \mid 0 \leq u < S_x ; 0 \leq v < S_y \},$$

je dimenzia konvulčného jadra, S_x a S_y sú rozmery konvulčnej mapy, I je súbor máp predchádzajúcej vrstvy, s ktorými je konvulčná mapa j prepojená. Nakoniec výstupná vrstva obsahuje 10 neurónov plne prepojených s predošlými priemerovacími mapami vrstvy H4, ktorá rozpoznáva 10 číslíc. Víťazný neurón (charakterizujúci rozpoznávanú číslicu) by mal byť rovný hodnote +1, pričom ďalšie by mali byť rovné hodnote -1. Celkovo má sieť 4635 jednotiek, 98442 prepojení, ale kvôli zdieľaniu váh iba 2578 parametrov je potrebné učiť.

3.2 Trénovanie konvulčných neurónových sietí

Error backpropagation algoritmus je najviac využívaným učiacim algoritmom pre KNS. Tento algoritmus je takmer identický s backpropagation algoritmom pre štandardný typ neurónových sietí. Rozdiel medzi týmito dvoma algoritmi je hlavne v zdieľaní váh v konvulčných a priemerovacích vrstvách. Taktiež chybová funkcia E bude minimalizovaná po každej tréningovej vzorke p :

$$E_p = \frac{1}{2} \| \mathbf{o}_p - \mathbf{t}_p \|^2 = \frac{1}{2} \sum_{k=1}^K (\mathbf{o}_{pk} - \mathbf{t}_{pk})^2$$

kde K je počet výstupných jednotiek, \mathbf{o}_{pk} je výstup neurónu k pre vzor p a \mathbf{t}_{pk} je príslušná cieľová hodnota \mathbf{t}_{pk} . Učiaci proces je iteratívna procedúra, kde každá úprava váhy je malým krokom k cieľu v porovnaní so smerovaním gradientu ∇E . Pre úpravu váh teda platí:

$$w_{ji}^{(l)} \leftarrow w_{ji}^{(l)} + \Delta w_{ji}^{(l)} = w_{ji}^{(l)} - \lambda \frac{\partial E_p}{\partial w_{ji}^{(l)}}$$

kde λ je učiaci pomer a $w_{ji}^{(l)}$ označuje váhu z neurónu i do neurónu j vo vrstve l .

Ak v NS je dodatočná skrytá vrstva obsahujúca jednoduché neuróny medzi poslednou priemerovacou vrstvou a výstupnou vrstvou, tak rovnica pre lokálnu chybu skrytého neurónu j vyzerá nasledovne:

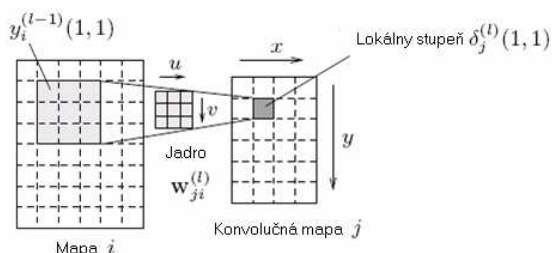
$$\delta_j^{(l)} = e_j^{(l)} \phi'(V_j^{(l)}) = \left(\sum_{k=1}^K \delta_j^{(l+1)} w_{kj}^{(l+1)} \right) \phi'(V_j^{(l)})$$

kde $e_j^{(l)}$ označuje chybu na neuróne j vo vrstve l spätne šírenú od neurónov z vrstvy $l+1$.

V prípade konvolučnej vrstvy je úprava príslušnej váhy $w_{ji}^{(l)}(u,v)$ zo znakovej mapy i do j na pozícii jadra (u,v) sumáciou cez všetky pozície (x,y) v znakovej mape.

$$\Delta w_{ji}^{(l)}(u,v) = -\lambda \sum_{(x,y)} \left(\delta_j^{(l)}(x,y) y_i^{(l-1)}(x+u, y+v) \right)$$

Obr. č. 4 znázorňuje Backpropagation s konvolučnými mapami a vzťah rôznych premenných i, j, u, v, x, y .



Obr. č. 4 Backpropagation s konvolučnými mapami

Odchýlka $b_j^{(l)}$ konvolučnej mapy j vo vrstve l je jednoducho upravovaná pripočítaním nasledujúcej hodnoty:

$$\Delta b_j^{(l)} = -\lambda \sum_{(x,y)} \delta_j^{(l)}(x,y)$$

S priemerovacími mapami je počítanie trochu odlišné. V tomto prípade je iba jedna váha $w_{ji}^{(l)}$ pre každé spojenie z mapy i do priemerovacej mapy j . Zvyčajne spojenie z konvolučnej do priemerovacej mapy je typu one-to-one, teda $i=j$. Vo všeobecnosti pre váhu $w_{ji}^{(l)}$ z mapy i vo vrstve $l-1$ do priemerovacej mapy j vo vrstve l so "sub-sampling" činiteľmi (alebo veľkosťou okna) S_x a S_y platí:

$$\Delta w_{ji}^{(l)} = -\lambda \sum_{(x,y)} \delta_j^{(l)}(x,y) \sum_{m=1}^{S_x} \sum_{n=1}^{S_y} y_i^{(l-1)}(x s_x + m, y s_y + n)$$

Obr. č. 5 zobrazuje vzťah príslušných premenných. Rovnica pre úpravu odchýlky je rovnaká ako pre konvolučné mapy.

Výpočet lokálneho chyby mapy j vo vrstve l závisí od typu nasledujúcej vrstvy $l+1$. Ak vrstva $l+1$ je neurónová vrstva, tak pre lokálnu chybu platí:

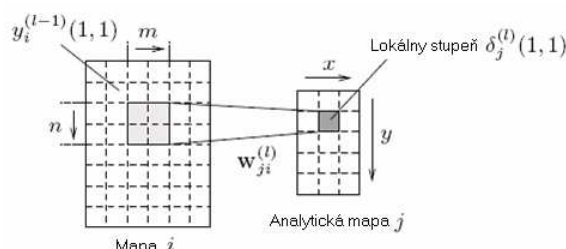
$$\delta_j^{(l)}(x,y) = \sum_{k=1}^K \sum_{(x,y)} \delta_k^{(l+1)} w_{kj}^{(l+1)}(x,y)$$

čo je zovšeobecnenie, ktoré berie do úvahy nielen spojenie jednotlivých neurónov, ale taktiež spojenie mapy s neurónmi $(x, y > I)$. K je počet neurónov vo vrstve $l+1$.

Ak nasledujúcou vrstvou je konvolučná vrstva, tak platí:

$$\delta_j^{(l)}(x,y) = \sum_{k \in K_c(u,v)} \sum_{(x,y)} \delta_k^{(l+1)}(x,y) w_{kj}^{(l+1)}(u,v)$$

kde K_c označuje súbor máp vo vrstve $l+1$ spojenú s konvolučnou mapou j vo vrstve l . Do úvahy treba zobrať to, že konvolučná maska nemôže presiahnuť hranicu vrstvy l .



Obr. č. 5 Backpropagation s priemerujúcimi mapami

V prípade, že vrstva $l+1$ je priemerovacou vrstvou, tak pre lokálny stupeň platí:

$$\delta_j^{(l)}(x,y) = \sum_{k \in K_s} \delta_k^{(l+1)} \left(\lfloor x/S_x \rfloor, \lfloor y/S_y \rfloor \right) w_{kj}^{(l+1)}$$

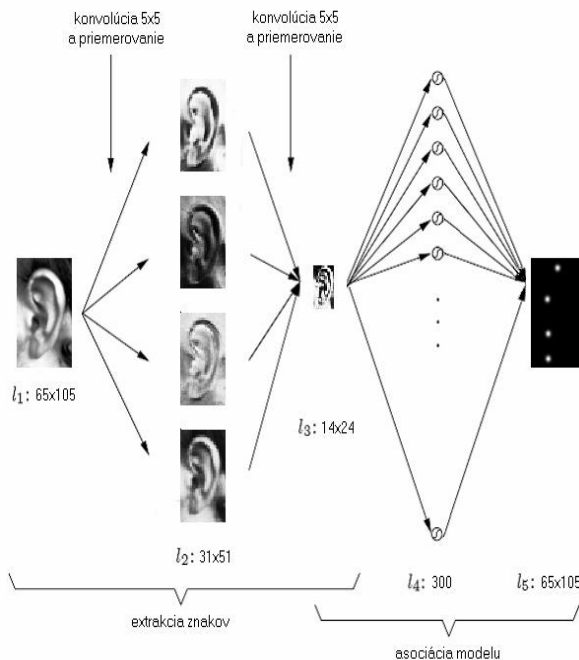
kde S_x, S_y sú činitele pre tvorbu menších vzoriek, K označuje súbor máp vo vrstve $l+1$ spojenými s priemerovacou mapou j vo vrstve l .

4 Popis systému na detekciu znakov ucha pomocou KNS

V tejto časti popíšeme systém na detekciu znakov ucha, ktorý je navrhnutý na presné lokalizovanie znakov ucha v čierno-bielych obrázkoch ucha. Popísaný systém vie nájsť znaky ucha v obrázkoch, ktoré boli fotené pri rôznej intenzite svetla, v rôznej kvalite a rozmeroch a s určitými rušivými momentmi. Tieto fotky boli potom upravované tak, aby boli vhodné na vstup do neurónovej siete. Neurónová sieť je založená na špecifickej architektúre konvolučných a hetero-asociačných neurónových vrstiev. Vrstvy sa skladajú z niekoľkých typov komponentov formujúcich sústavu špecializovaných transformácií obrázka, automaticky naučených z tréningovej vzorky uší a z ich vopred označených vybraných znakov ucha. Všetky komponenty sú sekvenčne prepojené a môžu byť trénované jednoduchým prezentovaním vstupného obrázka na vstupe a požadovaného výstupu, t. j. pozície znakov. Globálne obmedzenia sú automaticky učené a hneď použité v detekčnom procese. Po natrénovaní, systém na detekciu týchto znakov pôsobí ako sústava konvolučných a priemerovacích modulov, ktoré zo vstupného obrázka ucha budujú znakovú mapu, v ktorej sa znaky ucha dajú jednoducho nájsť.

4.1 Architektúra systému SDZU

Architektúra detektora znakov ucha je špecifický typ konvolučnej neurónovej siete (KNS) (viď Obr. č. 6) pozostávajúcej z 5 vrstiev, kde prvá vrstva je vstupná vrstva, druhá vrstva je konvolučná spojená s priemerovacou vrstvou, nasledujúca vrstva je konvolučná, a posledné 2 vrstvy sú štandardné neurónové vrstvy s dopredným šírením signálu. Cieľom systému je naučiť sa nelineárnu funkciu, ktorá transformuje vstupný obrázok ucha do požadovaného výstupu, t.j. znakovkej mapy, v ktorej sú vyznačené znaky ucha (viď Obr. č. 7).



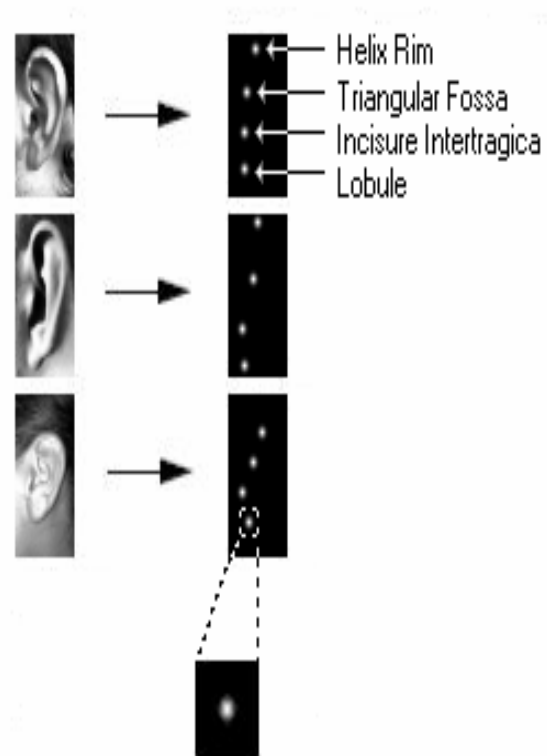
Obr. č. 6 Architektúra detektora SDZU znakov ucha

4.2 Popis jednotlivých vrstiev

Prvá vrstva l_1 (vstupná vrstva): Dostane na vstupe čierno-biely obrázok ucha rozmerov 105 x 65 pixlov, ktorý obsahuje hodnoty šedej normalizované medzi -1 a +1. Žiadne ďalšie spracovanie obrázka, ako zlepšenie kontrastu, redukcia rušivých momentov alebo aplikovanie nejakého filtrovania, nie je použité. Celkový počet neurónov v tejto vrstve je $105 \times 65 = 6825$. Váhy pre túto vrstvu nie sú definované.

Druhá vrstva l_2 : Je konvolučná vrstva spojená s priemerovacou. Pozostáva zo 4 znakových máp rozmerov 51 x 31. Každá znaková mapa má 5 x 5 tréované konvolučné jadro zo vstupnej vrstvy. Samotná metóda priemerovania [10] je vykonávaná posunom konvolučného jadra o každé

dva pixle po vstupnom obrázku. Celkový počet neurónov je $51 \times 31 \times 4 = 6324$ a definovaných váh je $(5 \times 5 + 1) \times 4 = 104$.



Obr. č. 7 Vstupné obrázky a požadované korešpondujúce výstupné znakové mapy (Helix Rim, Triangular Fossa, Incisure Intertragica, Lobule). Pre ilustračné účely obrázok v spodnej časti ukazuje zväčšenú oblasť jedného z požadovaných výstupných znakov.

Tretia vrstva l_3 : Je konvolučná vrstva spojená s priemerovacou vrstvou, rovnako ako predošlá vrstva. Pozostáva z jednej znakovkej mapy rozmerov 24 x 14 pixlov. Každý neurón z tejto vrstvy je spojený s 5 x 5 tréovaným konvolučným jadrom príslušnej znakovkej mapy, ktoré sú posúvané o dva pixle po všetkých 4 znakových mapách z predošlej vrstvy. Celkový počet neurónov je $24 \times 14 = 336$ a váh $(5 \times 5 + 1) \times 4 \times 1 = 104$.

Kým predošlé vrstvy sa správali hlavne ako vrstvy na extrakciu znakov, tak vrstvy l_4 a l_5 používajú lokálnu informáciu na vytvorenie globálneho modelu.

Štvrtá vrstva l_4 : Tvorí ju 300 neurónov plne prepojených so všetkými 336 neurónmi vo vrstve l_3 a sú určené na naučenie sa znakových modelov (alebo konštelácií) a aktivovanie cieľových pozícií vo výstupnej znakovkej mape. Celkový počet váh je $300 \times (336 + 1) = 101100$.

Táto časť neurónovej siete bola inšpirovaná Auto-Asociatívnymi neurónovými sieťami (AANS) (viď sekciu), ktoré sú trénované na asociovanie výstupu znakovej mapy v I_3 s požadovaným výstupom v I_5 . Preto dovoľíme nechať aktívované iba určité konštelácie znakov vo vrstve I_5 .

Piata vrstva I_5 (výstupná vrstva): Na výstupe je znaková mapa aktívovaná súborom naučených váh, ktoré dostanú globálny opis ucha vyprodukovaný vrstvou I_3 . Táto znaková mapa obsahuje všetky 4 body, ktoré majú byť nájdené a má rovnakú dimenziu ako obrázok na vstupnej vrstve, t.j. 105 x 65, čo znamená, že má celkovo 6825 neurónov. Táto vrstva je plne prepojená s neurónmi z predošlej vrstvy, čo znamená definovať celkovo 6825 x (300 + 1) = 2054325 váh.

Aktivačná funkcia: Všetky neuróny (samozrejme okrem neurónov prvej vrstvy) ako aktivačnú funkciu používajú tzv. propor-cionálny hyperbolický tangens [11]: $f(a) = A \tanh(Sa)$, kde A je amplitúda funkcie a S určuje jej východiskový bod. Funkcia f je nepárna, ohraničená horizontálnymi asymptotami $-A$ a $+A$. Táto funkcia bola vybratá, pretože pri symetrických funkciách sa vo všeobecnosti zdá, že konvergujú rýchlejšie, aj keď učenie sa môže spomaliť pri príliš malých váhach. Pre tento experiment boli použité hodnoty konštant $A = 1.7159$ a $S = 2/3$, pre ktoré boli splnené aj rovnosti $f(1) = 1$ a $f(-1) = -1$. Logické pri výbere týchto hodnôt je, že maximálne hodnoty, ktoré funkcia f môže nadobúdať, sú rovné približne 1, pričom sa interpretácia siete týmto iba zjednodušila. Dokonca aj maximum a minimum druhej derivácie funkcie f sú rovné $+1$ a -1 , čo nám zlepšuje konvergenciu a až do konca učiaceho procesu nemá žiadny dopad na celkový výsledok učenia sa siete.

4.3 Tréning detektora ušných znakov

Na tréningovanie neurónového detektora bol použitý modifikovaný backpropagation algoritmus pre konvolučné siete, ktorý zohľadňuje všetky úpravy navrhnuté siete a všetky možné zjednodušenia tréningu. Prípravu práce detektora znakov uší môžeme rozdeliť do troch fáz:

- Zbieranie dát a ich úprava
- Vytvorenie výstupných obrázkov
- Tréning

4.3.1 Zbieranie dát a ich úprava

Tréningové dáta, ako aj celá databáza obrázkov uší, bola získaná z rôznych zdrojov. Väčšina z nich bola fotená fotoaparátom na mobilnom telefóne Nokia N95 za rôznych podmienok, ale vo vzorke sa nachádzajú aj obrázky získané z internetu (vzorky 22 a 23) alebo fotené

regulárnym fotoaparátom (vzorky 8, 12 a 32). Pred použitím dát v experimente bolo potrebné vykonať niekoľko úprav. Okrem úprav ako centrovanie a zmenšovanie obrázkov na určitú veľkosť, ich bolo potrebné konvertovať do odtieňa šedej, tzn. že obrázok nebude po tomto konvertovaní farebný, ale čierno-biely. Po konvertovaní bola na obrázkoch použitá metóda "zarovnanie stĺpcového diagramu obrázka" (Histogram Equalization). Táto metóda sa používa v obrázkoch na lokálne zvýšenie kontrastu, čo znamená, že v miestach, kde je intenzita kontrastu menšia sa zvýši, bez nejakého ovplyvnenia globálneho kontrastu. Je využívaná najmä v prípadoch, kde v obrázkoch sú si hodnoty kontrastu veľmi blízke. Hodnoty kontrastu sú reprezentované stĺpcovým diagramom a pod jeho zarovnaním rozumieme, že hodnoty s najväčším výskytom sú rozptýlené po celom obrázku.

Metóda zarovnanie stĺpcového diagramu: V danom čierno-bielom obrázku, nech n_i je počet výskytov i -tého odtieňa šedej x_i . Pravdepodobnosť výskytu i -tého odtieňa šedej potom je:

$$p(x_i) = 2 \frac{n_i}{n} - 1, i \in 0, \dots, L-1$$

kde L je počet všetkých odtieňov šedej v obrázku, n je celkový počet pixlov v obrázku, a $p(x)$ predstavuje stĺpcový diagram obrázku, normalizovaný na $[-1, +1]$. Definujeme tiež c ako kumulatívnu distribučnú funkciu zodpovedajúcu p :

$$c(i) = \sum_{j=0}^i p(x_j)$$

tiež známou ako akumulovaný normalizovaný stĺpcový diagram obrázka. Teraz potrebujeme transformovať každý odtieň x v originálnom obrázku na odtieň y , čo urobíme pomocou lineárnej kumulatívnej pravdepodobnostnej funkcie y aplikovanej na všetky hodnoty. Transformácia potom bude definovaná nasledovne:

$$y_i = T(x_i) = c(i)$$

Treba poznamenať, že funkcia T zobrazuje odtiene do domény 0..1. Na to, aby sme získali hodnoty späť do ich originálnej domény, je potrebné na výsledné hodnoty z predchádzajúcej rovnice aplikovať jednoduchú transformáciu: $y_i = (2 * y_i - 1)(\max - \min) + \min$

4.3.2 Vytvorenie výstupných vzorov

Požadované výstupné mapy sú vytvorené tak, aby obsahovali hodnotu $+1$ na požadovanej pozícii znaku a -1 na ostatných pozíciách. Tieto hodnoty sa nebudú nadobúdať iba v okolí pozície znaku, kde budú hodnoty plynule klesať, z dôvodu lepšej konvergencie a robustnosti. Na vykonanie tohto plynulého klesania použijeme 2-dimenzionálnu Gaussovú funkciu centrovanú na pozíciu znaku

a normalizovanú medzi -1 a +1. Potom pre požadovanú znakovú mapu o a jej požadovaný znak na pozícii (u_x, u_y) funkcia vyzerá nasledovne:

$$o(x, y) = 2e^{-\frac{1}{2}\left(\frac{(x-u_x)^2}{\sigma_x^2} + \frac{(y-u_y)^2}{\sigma_y^2}\right)} - 1$$

V experimentoch boli hodnoty premenných σ_x a σ_y nastavené na 2.

4.3.3 Trénovanie

V tréningovej fáze bol použitý štandardný online backpropagation algoritmus s potrebnými zmenami. Pri každej iterácii bol na vstup systému prezentovaný postupne každý obrázok z tréningových dát a podľa toho boli upravované aj váhy (online tréning). Pre vyhnutie sa problémom zo zlým nastavením po každej iterácii je spustená fáza validácie na validačných dátach. Znamená to, že sa hľadá minimálna chyba na validačných dátach, tak aby sa dosiahol najlepší výsledok, po ktorého nájdení je príslušná váhová konfigurácia uložená. Do úvahy prichádzali tieto kritéria na minimalizovanie chyby:

- **Kritérium 1.** Chyba hľadaná medzi hodnotami výstupnej mapy $y_k(i, j)$ a príslušnými hodnotami požadovanej výstupnej mapy $O_k(i, j)$, t.j. chyba počítaná neurón po neuróne:

$$E_{MSE} = \frac{1}{FNM} \sqrt{\sum_{k=1}^F \sum_{i=1}^N \sum_{j=1}^M (y_k(i, j) - o_k(i, j))^2}$$

kde F je číslo výstupnej znakovkej mapy a N x M jej dimenzia.

- **Kritérium 2.** Určitý typ vzdialenosti medzi nájdenými výstupnými pozíciami znaku (i_{yk}, j_{yk}) a požadovanými výstupnými pozíciami znaku (i_{ok}, j_{ok}) :

$$E_{dist} = \frac{1}{F} \sqrt{\sum_{k=1}^F \left((i_{yk} - i_{ok})^2 + (j_{yk} - j_{ok})^2 \right)}$$

kde

$$(i_{yk}, j_{yk}) = \arg \max_{i, j} y_k(i, j) \text{ a}$$

$$(i_{ok}, j_{ok}) = \arg \max_{i, j} o_k(i, j)$$

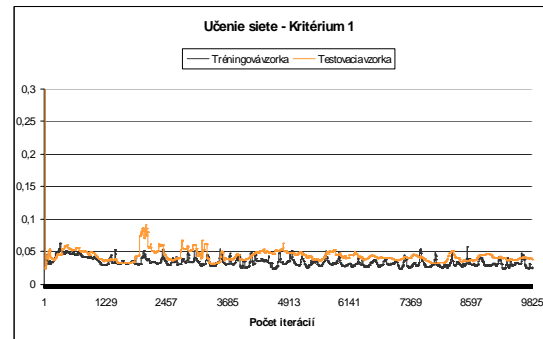
- **Kritérium 3.** Chyba hľadaná v maličkom štvorci v okolí požadovaných výstupov. Ak je hodnota odtieňa sivej aspoň troch a viac pixlov, ktoré sú vzdialené najviac jeden pixel od požadovaného výstupu, väčšia ako 172, tak je znak posúdený ako nájdený.

V systéme SDZU bola využitá kombinácia prvého a tretieho kritéria.

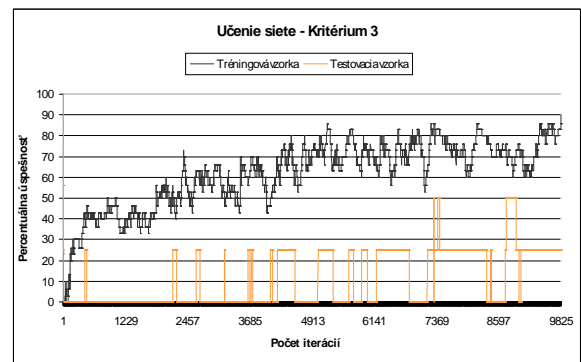
5 Vyhodnotenie dosiahnutých výsledkov

Experiment bol vykonaný na vzorke 34 obrázkov ucha. SDZU bol tréňovaný na 30-tich obrázkoch z tejto vzorky a zvyšné štyri slúžili ako testovacie. Hlavnými kritériami pre správne nájdenie identifikačných znakov bolo prvé a tretie kritérium na minimalizovanie chyby, ktoré sú popísané vyššie. Keď boli tieto kritéria splnené na určité vopred stanovené hodnoty, tak bolo vykonané ešte manuálne porovnanie výstupov zo siete s tými, ktoré boli označené.

Pri tréňovaní siete bolo potrebné nájsť čo najlepšiu hodnotu učiacich pomerov pre jednotlivé vrstvy. Nakoniec počiatočná hodnota učiaceho pomeru λ pre tretiu a štvrtú vrstvu bola určená na 0.00001, ktorá sa po každej šiestej iterácii prenášobila hodnotou 0.794183335. Násobený bol dovtedy, kým nedosiahol svoje minimum, ktoré bolo nastavené na hodnotu 0.000000005. Učiace pomery pre prvú a druhú vrstvu boli vyrátavané z hodnoty predošlého učiaceho pomeru a to nasledovným spôsobom $\lambda^1 = \lambda / 50000$ a $\lambda^2 = \lambda / 9000$. Priebeh tréňovania siete podľa jednotlivých kritérií je zobrazený na nasledujúcich dvoch obrázkoch Obr. č. 8 a Obr. č. 9.



Obr. č. 8 Trénovanie siete pri zohľadnení kritéria 1.



Obr. č. 9 Trénovanie siete pri zohľadnení kritéria 3.

Sieť bola považovaná za naučenú, keď na tréningovej vzorke chyba podľa kritéria 1 dosiahla hodnotu pod 0.03 a zároveň hodnota podľa kritéria 3 mala úspešnosť väčšiu ako 89%. Obidve kritéria sieť splnila po 9824-tej iterácii, pričom boli dosiahnuté nasledujúce hodnoty:

Tréningová vzorka

Kritérium 1 = 0.0261015 a Kritérium 3 = 90%

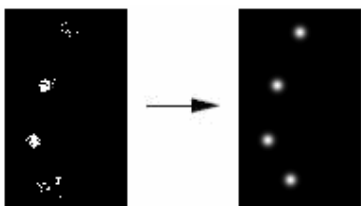
Testovacia vzorka

Kritérium 1 = 0.0383563 a Kritérium 3 = 25%

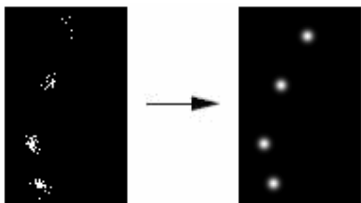
Pri analýze jednotlivých výsledkov sme zistili, že pre niektoré vzorky neboli nájdené všetky štyri identifikačné body ucha celkom presne.

V tréningovej vzorke bolo 30 vzoriek a problémy vznikli pri troch vzorkách. Naučenej sieti nepodarilo nájsť tieto body (číslované zhora smerom nadol):

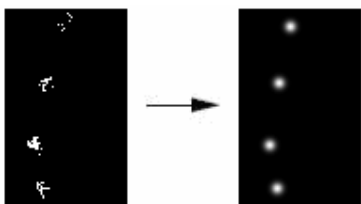
Vo vzorke 5 prvý a štvrtý bod,



vo vzorke 9 prvý a druhý bod

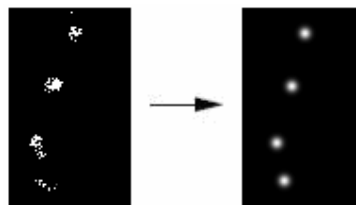


a vo vzorke 29 prvý bod.

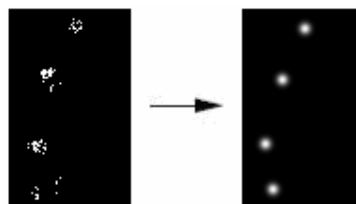


V testovacej vzorke, ktorá mala štyri vzorky, ukážeme situáciu pre všetky štyri. V tomto prípade sme pozorovali väčší rozptyl pixlov u niektorých bodov (číslované zhora smerom nadol):

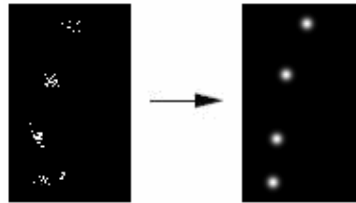
Vo vzorke 31 tretí a štvrtý bod,



vo vzorke 32 štvrtý bod



a vo vzorke 34 tretí a štvrtý bod.



Po naučení siete bola ešte zisťovaná presná pozícia odhalených znakov a tá bola porovnávaná s požadovanou pozíciou znakov. Priemerná vzdialenosť medzi týmito dvoma znakmi na jeden obrázok na tréningovej vzorke bola 1.85 pixla a na testovacej 3.85 pixla.

6 Záver

V príspevku je popísaná problematika detekcie ušných znakov (boli vybraté 4 znaky) v obrázkoch ucha, čo je považované za jeden z najdôležitejších krokov pri vývoji systémov na klasifikáciu, resp. na analýzu ucha. Bol vytvorený špeciálny typ neurónovej siete, nazývaný konvolučná neurónová sieť. Táto sieť bola využitá v experimente, kde boli hľadané štyri určité body na uchu. Sú tu prezentované výsledky z experimentu, z ktorých je

zrejme, že dané body na obrázkoch ucha je možné nájsť s určitou požadovanou presnosťou a chybovosťou. Niektoré body aj napriek vysokej úspešnosti nájdenia bodov a nízkej chybovosti neboli nájdené. Príčin, prečo daná sieť niektoré body nemohla nájsť môže byť viacej (napr. malá tréningová vzorka, zvolenie hodnôt učiacich pomerov, atď.).

Ďalším pokračovaním tohto experimentu by mohla byť extrakcia abstraktnejších znakov ucha, ako napr. vzdialenosti medzi jednotlivými časťami ucha, krivky alebo iné geometrické parametre, na ktorých by identifikácia mohla byť založená. Ďalšie parametre by mohli upresniť identifikáciu, resp. klasifikáciu uši.

Literatúra

- [1] Fields, C., Hugh, C.F., Warren, C.P. and Zimberoff, M.: The ear of the newborn as an identification constant. *J. Obst. Gynaec.*, 16: 98-101 (1960).
- [2] Imhofer, R.: Die Bedeutung der Ohrmuschel für die Feststellung der Identität. *Archiv für die Kriminologie*, 26: 150-163 (1906).
- [3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner: Gradient-Based Learning Applied to Document Recognition, *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.
- [4] A. Duffner: Face Image Analysis With Convolutional Neural Networks, Dizertačná práca, Fakultät für Angewandte Wissenschaften an der Albert-Ludwigs-Universität Freiburg im Breisgau, 2007
- [5] A. Iannarelli, Ear Identification. Forensic Identification Series. Paramount Publishing Company, Fremont, California
- [6] B. Victor, K. W. Bowyer, and S. Sarkar. An evaluation of face and ear biometrics. In *Proc. ICPR*, strany 429-432, 2002.
- [7] K. Chang, K. Bowyer, and V. Barnabas. Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 25:1160–1165, 2003.
- [8] Ping Yan, Kevin W. Bowyer: Empirical Evaluation of Advanced Ear Biometrics, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* Washington, DC, USA, 2005.
- [9] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel. Handwritten digit recognition with a back-propagation network. In David Touretzky, editor, *Advances in Neural Information Processing Systems 2*, pages 396–404. Morgan Kaufman, Denver, CO, 1990
- [10] Dr. P. Simard: Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis
- [11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner: Gradient-Based Learning Applied to Document Recognition, *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.
- [12] T. Yuizono, Y. Wang, K. Satoh, and S. Nakayama: Study on individual recognition for ear images by using genetic local search. In *Proc. of the 2002 Congress on Evolutionary Computation*, 2002, p. 237-242.
- [13] M. Choraš. Ear biometrics based on geometrical feature extraction. *Electronic Letters on Computer Vision and Image Analysis*, 5:84–95, 2005
- [14] H. Chen and B. Bhanu. Contour matching for 3D ear recognition. In *Seventh IEEE Workshop on Application of Computer Vision*, strany 123-128, 2005
- [15] D. J. Hurley. Force Field Feature Extraction for Ear Biometrics. PhD thesis, Electronics and Computer Science, University of Southampton, 2001.
- [16] B. Moreno, A. Sanchez, and J. Velez, "On the Use of Outer Ear Images for Personal Identification in Security Applications," *Proc. IEEE Int'l Carnaham Conf. Security Technology*, pp. 469-476, 1999.
- [17] T. Yuizono, Y. Wang, K. Satoh, and S. Nakayama. Study on individual recognition for ear images by using genetic local search. In *Proceedings of the 2002 Congress on Evolutionary Computation*, strany 237–242, 2002.