

Az enyhe kognitív zavar és korai Alzheimer-kór automatikus azonosítása spontán beszédből akusztikus jellemzők segítségével

Gosztolya Gábor ¹, Hoffmann Ildikó ^{2,3}, Tóth László ⁴,
Vincze Veronika ^{1,5}, Pákáski Magdolna ⁶, Kálmán János ⁶

¹MTA-SZTE Mesterséges Intelligencia Kutatócsoport, Szeged

²Szegedi Tudományegyetem, Magyar Nyelvészeti Tanszék

³MTA Nyelvtudományi Intézet, Budapest

⁴Szegedi Tudományegyetem, Informatikai Intézet

⁵Szegedi Tudományegyetem, Általános Nyelvészeti Tanszék

⁶Szegedi Tudományegyetem, Pszichiátriai Klinika
{ ggabor, tothl, vinczev } @ inf.u-szeged.hu

Kivonat Az enyhe kognitív zavar (EKZ) olyan tünetegyüttes, melynek fontos szerepe van néhány demenciatípus, például egy később lehetségesen kialakuló Alzheimer-kór korai megjósolásában, és így a kezelés minél korábbi elkezdésében. Korábbi kutatásainkban megmutattuk, hogy a korai vagy enyhe Alzheimer-kór (eAK) hatékonyan felismerhető a vizsgálati személy spontán beszéde alapján a megfelelő beszédjellemzők kinyerése révén. Ebben a cikkben a korábban bemutatott, beszédfelismerésen alapuló automatikus jellemzőkinyerést alkalmazzuk az eAK-ban szenvedő és az egészséges (kontroll) beszélők megkülönböztetése mellett az EKZ-val élők felismerésére is. A kísérleti eredmények alapján a kontroll csoport és az eAK beszélők megkülönböztetése történik a legnagyobb hatékonysággal, ugyanakkor az EKZ-val élő betegek felismerése is hatékony: a páronkénti pontosságértékek 74% és 84% közé esnek.

Kulcsszavak: beszédfelismerés, beszédfeldolgozás, demencia, enyhe kognitív zavar, korai Alzheimer-kór

1. Bevezetés

A demenciában szenvedő személyek száma az életkor kitolódása miatt az elkövetkező néhány évtizedben várhatóan megsokszorozódik [1]. Mindeközben a demencia korai felismerése nem egyszerű, és a diagnosztizált esetek aránya is alacsony. A demencia egyik leggyakoribb típusa az Alzheimer-kór (AK), melynek kezelése a betegség korai szakaszában a leghatékonyabb. A korai felismerés fontos a későbbi kezelés szempontjából, hiszen így lassítható a betegség előrehaladása és javítható a beteg életminősége.

A demetálódás folyamata már negyvenéves kor körül elkezdődhet az úgynevezett enyhe kognitív zavar (EKZ) tünetegyüttesel. Az enyhe kognitív zavar a

különböző mentális funkciók egy megelőző magasabb szintjéhez képest tapasztalható kezdődő, enyhe változás. Ez megjelenhet a rövidtávú emlékezet kapacitásának finom korlátozódásában, a szókeresés gyakoriságának enyhe növekedésében vagy figyelmi zavarokban. Az EKZ az egészséges és a demens állapot közötti skálán jelenik meg; előjele lehet egy később kialakuló demenciának. A demencia ugyanakkor folytonosan meglévő kognitív deficitet jelent a mentális funkciók legalább három területén, mely az emlékezetet, a nyelvet, a téri-vizuális készséget, az absztrakt gondolkodást, a számolást, az ítéletképességet, az érzelmi állapotot vagy a személyiség megváltozását érintheti [2,3,4].

A klinikai pszichiátria egyik legnehezebb feladata az EKZ diagnosztizálása. Habár a nemzetközi szakirodalom és a klinikai gyakorlat számos korai felismerést célzó diagnosztikai tesztet ismer — pl. Mini Mentál Teszt (MMT, [5]), Órarájzoló Teszt (Clock Drawing Test, CDT [6,7]), ADAS-Cog ([8]) —, ezek vagy túl időigényesek, vagy nem elég érzékenyek ahhoz, hogy felismerjék azokat a finom változásokat, melyek az EKZ-ban megjelennek, így eredményeik is gyakran egymásnak ellentmondóak. Léteznek ugyanakkor olyan pszicholingvisztikai és neurolingvisztikai tesztek, mint például a szótalálás vagy a történetfelidézés, melyek jelentősen hatékonyabbnak bizonyulnak az EKZ felismerésére, ugyanakkor ezen tesztek eredményei félre is diagnosztizálhatnak [9].

Újabb tanulmányok rámutatnak arra, hogy az AK szubklinikus szakaszában, azaz az EKZ-ban már tapasztalhatók nyelvi változások. Ezek a változások érinthetik a beszéd temporális jellemzőit [9,10,11,12,13,14,15], a mentális lexikon elemeinek lehívását [16,17,18,19,20,15] és a verbális fluenciát [21,9,18,14] is.

Jelen cikkünkben a beszéd temporális paramétereinek elemzésével foglalkozunk, melynek segítségével olyan beszéd közbeni kognitív folyamatokat vizsgálhatunk, mint a nyelv és a memória közti kapcsolatok. Végső célunk egy automatikus szoftveres alkalmazás kifejlesztése, mely alkalmas a potenciális páciensek előszűrésére, elősegítve az EKZ-ban és az enyhe AK-ban szenvedők korai diagnosztizálását. Erre egy, korábbi tanulmányainkban ([11,22,23,24]) már bemutatott automatikus kétlépéses eljárást alkalmazunk. Jelen cikkünkben az enyhe vagy korai AK (eAK) felismerése mellett célunk az EKZ detektálása is, valamint a két betegcsoport egymástól való elkülönítése.

2. A felhasznált adatbázis

A hangfelvételeket a Szegedi Tudományegyetem Pszichiátriai Klinikáján rögzítettük az EKZ és korai AK által leginkább veszélyeztetett korosztálytól. A kutatást a Szegedi Tudományegyetem Etikai Bizottsága hagyta jóvá, és azt a Helsinki Nyilatkozat betartásával végeztük. Három beszélőcsoporttól rögzítettünk felvételeket: EKZ-val és korai AK-val élők, valamint olyan kontroll beszélőktől, akik a felvétel időpontjában nem szenvedtek semmilyen mentális betegségben. A három csoportot korban, nemben és iskolázottságban illesztettük. A kontroll csoport alkohol vagy mentális funkciókat befolyásoló gyógyszerhatás alatt nem állt, sem látási, sem hallási deficittel nem rendelkezett. Kizártuk azokat, akik korábban fejsérülést szenvedtek, és azokat is, akik depressziós tüneteket mutattak.

	Beszélőcsoportok			Stat.
	Kontroll	EKZ	eAK	p
Életkor	$70,72 \pm 5,004$	$72,4 \pm 3,594$	$73,96 \pm 6,846$	$p = 0,105$
Iskolázottság (év)	$12,08 \pm 2,326$	$10,84 \pm 2,304$	$10,76 \pm 2,818$	$p = 0,118$
MMT érték	$29,24 \pm 0,523$	$27,16 \pm 0,898$	$23,92 \pm 2,488$	$p < 0,001$
CDT érték	$8,88 \pm 2,007$	$6,44 \pm 3,429$	$5,88 \pm 3,244$	$p = 0,001$
Adas-COG érték	$8,575 \pm 2,374$	$12,044 \pm 3,205$	$18,675 \pm 5,818$	$p < 0,001$

1. táblázat. A vizsgálati csoportok demográfiai adatai.

Az EKZ és eAK személyek orvosilag diagnosztizáltak voltak, velük CT, MRI és általános kognitív állapotot felmérő tesztek készültek (MMT, Órarajzolósi teszt, ADAS-Cog).

Mivel mind saját korábbi kutatásaink ([11,22,23,24]), mind más kutatócsoportok által végzett kutatások (például [17,9,13]) azt erősítették meg, hogy az EKZ és az eAK jobban befolyásolja a beszélő spontán beszédét, mint az előre eltervezettet, célunk spontán beszéd rögzítése volt. Emiatt jelen kísérletünkben a korábbi munkánkban bemutatott módon készítettük pácienseinket spontán beszédre, az alábbi forgatókönyv szerint. Miután megnézték egy egyperces animációs filmet, tesztalanyainkat megkértük, hogy meséljék el a filmben látottakat (*azonnali felidézés*). Ezután a vizsgálati személyeket megkértük, hogy meséljék el a tegnapi napjukat (*spontán beszéd*). Végül egy második egyperces filmet vetítettünk le, majd a teszt személyeknek egyperces várakozás után kellett elmesélniük annak tartalmát (*késleltetett felidézés*). (A részletekért ld. Hoffmann és mtsai cikkét [11].)

EKZ és korai AK vizsgálati személyeink hangfelvételeit tartalmazó adatbázisunk folyamatosan gyarapszik; jelenleg több, mint 150 beszélő spontán beszédét tartalmazza. Ugyanakkor különböző okokból (pl. ellentmondásos diagnózis vagy rossz hangminőség) néhány személy felvételeit ki kellett zárunk a további vizsgálatokból. Ezenfelül, mivel a három beszélői csoportot illeszteni kívántuk életkor, nem és iskolázottság szerint, további hangfelvételeket kellett kihagynunk jelen tanulmányunkból. Végül minden beszélői csoportból 25-25-25 személy anyagait használtuk, így kísérleteinket összesen 75 beszélő 225 hangfelvételén végeztük. Az 1. táblázat tartalmazza a kísérletben vizsgált alanyok néhány fontosabb adatát.

3. EKZ felismerésére használt beszédjellemzők

A spontán beszéd időviszonyainak elemzése különösen érzékeny neuropszichológiai módszernek bizonyult a kognitív folyamatok (mint például a beszéd) vizsgálatára [11]. A spontán beszéd temporális paramétereinek vizsgálata nagy pontossággal teszi mérhetővé a beszélő személy beszéd- és nyelvi képességeit [25,26].

Több tanulmány (pl. [21,27]) is megmutatta, hogy a demencia megléte növeli a hezitálás mennyiségét és időtartamát a spontán beszédben. Korábbi munká-

- (1) Artikulációs tempó (hezitációk nélkül számított másodpercenkénti beszédhang-szám).
- (2) Beszédtempó (hezitációkkal együtt számított másodpercenkénti beszédhang-szám).
- (3) A felvétel hossza, ezredmásodpercben mérve.
- (4-5) Kitöltött, illetve kitöltetlen szünetek összhossza.
- (6-7) Kitöltött, illetve kitöltetlen szünetek száma.
- (8) Hezitációs ráta (a kitöltött és kitöltetlen szünetek összhosszának és a felvétel hosszának aránya).

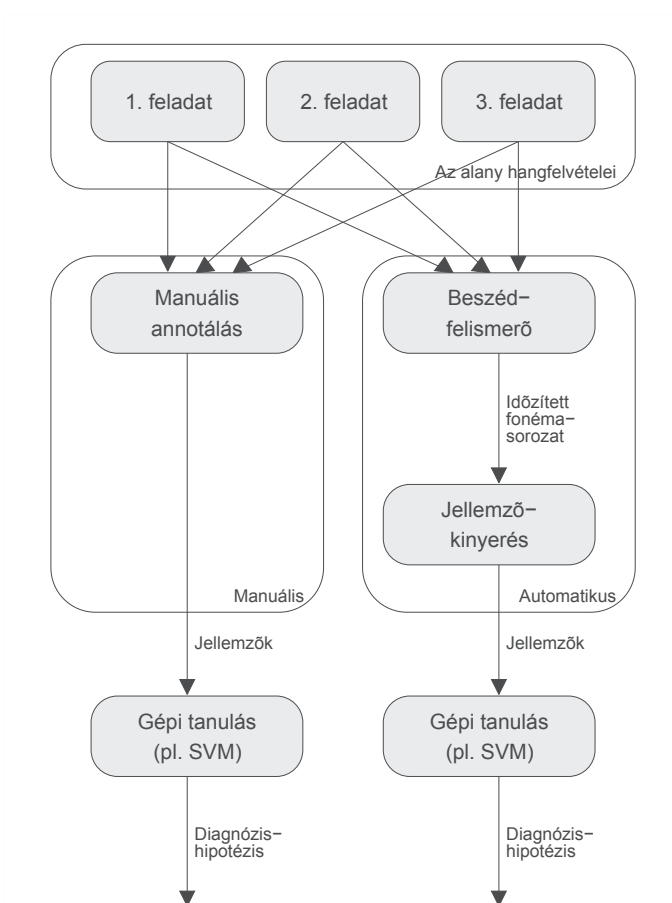
2. táblázat. Az alapvető akusztikus jellemzők, Hoffmann és mtsai [11] nyomán.

inkat (pl. [11]) követve, első lépésként a következő nyolc akusztikus jellemzőt határoztuk meg a manuális méréseinkhez, melyeket a Praat program segítségével végeztünk el: artikulációs tempó (1. jellemző), beszédtempó (2. jellemző), a jelidő hossza (3. jellemző), kitöltött, illetve kitöltetlen szünetek összhossza (4-5. jellemzők), kitöltött, illetve kitöltetlen szünetek száma (6-7. jellemzők), hezitációs ráta, azaz az összes szünet aránya a jelidőben (8. jellemző). Hezitáción a 30 ms hosszúságúnál nagyobb szünetet értjük [28]. A nyolc alapvető akusztikus paramétert és kiszámításuk módját a 2. táblázat foglalja össze.

3.1. A jellemzők automatikus kinyerése beszédfelismerési technikákkal

A fenti jellemzők manuális kiszámítása meglehetősen munkaigényes, ráadásul hozzáértő annotátorokat igényel. Ezért kíváncsiak a jellemzők kinyerésének automatizálása, melyet mi beszédfelismerési technikák használatával oldottunk meg. A jellemzők egy része ugyan egyszerű jelfeldolgozási eszközökkel is meghatározható (pl. beszéd/csend részek elkülönítése), azonban a beszédtempó és egyéb más, a beszédhangok hosszán alapuló jellemzők kinyerésére ezek alkalmatlanok. Ezen felül a kitöltött szüneteket is képtelenek megkülönböztetni az alany beszédétől.

A fentiek miatt egy beszédfelismerő rendszert tanítottunk be egy spontán beszédet tartalmazó adatbázisra. Vegyük észre, hogy a korábban felsorolt jellemzők kiszámításához általában nem szükséges az egyes beszédhangok *megkülönböztetése*, csupán azok *megszámolása*, illetve hosszának megmérése. Mivel a beszédhangok téves azonosítása általában nem okoz gondot, ezért egy egyszerűsített beszédfelismerő modellt használtunk, amely nagy pontossággal képes megtalálni és azonosítani a beszédjelenben szereplő beszédhangokat [29]. A kitöltött szüneteket önálló „fonémának” tekintettük, így a beszédfelismerő rendszer kimenete minden felvétele egy időzített beszédhang-sorozat. Ennek alapján a korábban javasolt jellemzőket már automatikusan is ki tudjuk számítani. A beszédjellemzők alapján a beszélőket valamilyen gépi tanulási eljárás (pl. SVM, véletlen erdő) segítségével, automatikusan különítjük el (ld. 1. ábra).



1. ábra: Az alkalmazott jellemzőkinyerési és osztályozási megközelítés manuális (bal oldal) és automatikus (jobb oldal) jellemzőkinyerés esetén.

Korábbi tanulmányaink ([22,23,24]) kísérleti eredményei igazolták, hogy a fenti automatikus jellemzőkinyerés alkalmazható az EKZ-ban szenvedő beszélők nagypontosságú megkülönböztetésére az egészséges kontroll alanyoktól. Jelen munkánkban ugyanezt a jellemzőkészletet és jellemzőkinyerési megközelítést használjuk, viszont most mindhárom beszélőcsoportot (EKZ, korai AK és kontroll) el szeretnénk különíteni egymástól.

3.2. Kibővített jellemzőkészlet

A fenti (ld. 2. táblázat) jellemzők manuális számításokra lettek tervezve. Ugyanakkor a 3.1. fejezetben ismertetett automatikus beszédfelismerésen alapuló módszer a felvételhez első lépésben egy időzített fonémasorozatot határoz meg, majd

- (1) Az adott fonéma előfordulásainak száma osztva a felvétel összfonémaszámával.
- (2) Az adott fonéma összhossza osztva a felvétel teljes hosszával.
- (3) Az adott fonéma előfordulásainak átlaghossza.
- (4) Az adott fonéma előfordulásainak hosszainak szórása.

3. táblázat. A négy statisztikai fonémajellemző, Tóth és mtsai ([22]) nyomán.

abból számítja ki a felvételenként nyolc beszédjellemzőt. A folyamat legidőigényesebb lépése nyilvánvalóan a beszédfelismerő rendszer alkalmazása, így a beszédjellemzők tényleges meghatározása csak elhanyagolható arányban járul hozzá a teljes elemzés futási idejéhez. Ezen a ponton könnyen bővíthetjük a számolt beszédjellemzők listáját további jellemzőkkel, amennyiben azok az időzített fonémamasorozatból számolhatóak. Ez a felismerés vezetett minket arra, hogy további, az EKZ (és remélhetőleg a korai AK) felismerését segítő jellemzőket adjunk a meglévő jellemzőhalmazhoz (*kiterjesztett jellemzőkészlet*).

Első lépésben megtartottuk a 2. táblázatban felsorolt jellemzőket, bár a (6)-os és (7)-es jellemzőket kicsit módosítottuk: ahelyett, hogy a kitöltött és kitöltetlen szünetek nyers darabszámára hagyatkoztunk volna, normalizált értékeket számoltunk úgy, hogy a szünetek számát elosztottuk az adott hangfelvétel fonémaszámával. A következő lépésben, mivel már rendelkezésünkre állt a kitöltött és kitöltetlen szünetek minden egyes előfordulásának hossza, a beszédjellemzők közé felvettük ezen hosszok átlagát és szórását is.

Észrevettük továbbá, hogy a beszédfelismerő rendszer gyakran keverte össze a kitöltött szüneteket bizonyos fonémákkal; például az „öö” hezitáció a fonetikai szintű kimenetben többször egyszerű „ö” fonémaként jelent meg, míg a „hmm” hezitáció rövid vagy hosszú „m”-ként. Ez alapján azt feltételeztük, hogy ezen fonémák számának, illetve összhosszának növekedése a fonetikai szintű beszédfelismerő-kimenetben valószínűleg a kitöltött szünetek számának gyarapodását jelzi, melyeket az akusztikus modell rosszul kategorizált. Így akusztikus jellemzőkészletünket olyan jellemzőkkel bővítettük ki, melyek ezen fonémák adott hangfelvételen belüli gyakoriságát írják le: az „m”, „n” és „ö” fonémák előfordulásából négy-négy jellemzőt számoltunk (ld. 3. táblázat). Ezzel a kiterjesztéssel a három felvételre összesen 81 jellemzőt kaptunk, melyre kísérleteink során mint a *kibővített* jellemzőkészletre fogunk hivatkozni.

Az EKZ és korai AK kockázatai különböznek férfiak és nők esetében, és a korrallal is változnak [30]. Mivel mind a nem, mind az életkor rendelkezésünkre állt a kísérletben részt vevő összes alany esetében, ezeket a jellemzőket és az iskolázottságot (iskolában elvégzett évek számát) is hozzávettük a jellemzőkészlethez, így alap jellemzőkészletünk 27, míg kiterjesztett jellemzőhalmazunk végül 84 jellemzőből állt. Habár mind a nem, mind az életkor elég megbízhatóan meghatározható a beszédhangból (ld. pl. [31,32]), a tervezett alkalmazásban egyszerűen meg is kérhetjük az alanyt, hogy adja meg ezeket az adatokat.

4. Kísérletek és eredmények

Az automatikus beszédfelismerő rendszert a BEA Spontánbeszéd-adatbázis [33] kb. hétórányi beszédatására tanítottuk be. Mivel fontos volt, hogy a kitöltött szüneteket képesek legyünk azonosítani, az elérhető annotációt módosítanunk kellett, hogy tartalmazzon bizonyos, a spontán beszédben előforduló elemeket (pl. kitöltött szüneteket, be- és kilégzéseket, nevetést, köhögést). Fonémaszintű felismerést végeztünk, a kitöltött szüneteket egy speciális fonémának tekintve. Akusztikus modellként egy standard előrecsatolt (feed-forward) neurális hálót alkalmaztunk három rejtett réteggel, rétegenként ezer ún. rectified aktivációs függvényt használó neuronnal. Nyelvi modellünk egy egyszerű fonémabigram volt, mely tartalmazta a fent felsorolt nemverbális hangjelenségeket (kitöltött szünetek, köhögések stb.).

A beszédfelismerő rendszer kimenete egy időzített, fonetikai szintű kimenet volt, mely tartalmazta mind a kitöltött, mind a kitöltetlen szüneteket. Ebből a fonetikai szintű átiratból már mind az alapvető, mind a kibővített jellemzőkészletet könnyen ki lehetett számítani.

4.1. Beszélőosztályozás

Gépi tanulási szemszögből egy 75 példából (most: alanyból) álló adatbázis nagyon kicsinek számít. Azonban a diagnosztizált EKZ és korai AK alanyok száma erősen véges, spontán beszédük rögzítése fix protokoll mellett pedig nagyon munkaigényes. Az általunk ismert többi, hasonló területet vizsgáló tanulmány (pl. [13,14,34,9]) mind 100-nál kevesebb beszélőt vizsgált.

Mivel ilyen kisméretű adatbázissal dolgoztunk, az alanyokat nem osztottuk tovább külön tanító- és tesztalmazra, hanem ötszörös keresztvalidálást (cross-validation, CV) alkalmaztunk. Hogy minden esetben minden beszélőcsoportból ugyanannyi jusson mind az aktuális tanító-, mind az aktuális tesztalmazba: mindig 60 beszélő felvételein tanítottunk, melyből 20 EKZ-ban, 20 enyhe AK-ban szenvedett, míg 20 a kontrollcsoportba tartozott. Ezután a betanított gépi tanulási modellt kiértékeltek a fennmaradó 15 beszélőre. Ezt ötször megismételve az összes beszélőre kaptunk egy diagnózis hipotézist, melyeket egyetlen végső pontosságértékben összegeztünk.

Osztályozásra szupportvektor-gépet (Support-Vector Machines, SVM [35]) alkalmaztunk lineáris kernellel; a Weka ingyenesen elérhető programcsomagban [36] található *SMO* implementációt használtuk.

4.2. Kiértékelés

A kiértékelési metrika kiválasztása nem egyértelmű ezen a feladaton. Használhatjuk a hagyományos osztályozási pontosságot, hiszen jelen esetben az osztályeloszlás tökéletesen kiegyensúlyozott. Ugyanakkor amellett, hogy azt számoljuk ki, mennyire pontosan sikerült azonosítanunk a három beszélőcsoport tagjait, az adott feladatot detektálási feladatnak is tekinthetjük, melyben a kérdés az,

Jellemző- kinyerés	Jellemző- készlet	3-oszt.	2-osztályos				
		Pont.	Pont.	UAR	Prec.	Fedés	F_1
Manuális	Alapvető	61,3%	76,0%	74,0%	83,3%	80,0%	81,6%
Automatikus	Alapvető	61,3%	70,7%	67,0%	78,0%	78,0%	78,0%
	Kibővített	62,7%	74,7%	73,0%	83,0%	78,0%	80,4%

4. táblázat. A különböző akusztikus jellemzőkészletek használatával elért pontosságértékek háromosztályos modellt tanítva.

hogy az adott beszélőnek van-e *bármilyen* mentális betegsége. Ezt a gyakorlatban úgy a legegyszerűbb megoldani, hogy az EKZ és a korai AK csoportokat egyesítjük. Mivel ekkor az osztályok példaszámai kiegyensúlyozatlanná válnak (hiszen 25 kontroll és 50 valamilyen fajta demenciában szenvedő alany van), a (kétoosztályos) osztályozási pontosság mellett az átlagos fedést (*Unweighted Average Recall*, *UAR*) is feltüntetjük. Emellett a szokásos információ-visszakeresési metrikákat is kiszámítottuk: pontosságot (*precision*, *Prec.*), fedést (*recall*, *Rec.*) és F -értéket (*F-measure* vagy F_1). Az SVM C meta-paraméterét úgy állítottuk be, hogy a legmagasabb (háromosztályos) osztályozási pontosságot kapjuk; azonos pontosságértékek esetén a legmagasabb F -értékkel rendelkező konfigurációt választottuk.

4.3. Eredmények

A 4. táblázat mutatja az osztályozási eredményeket a különböző akusztikus jellemzőkészletek használata mellett. Első pillantásra a háromosztályos pontosságértékek alacsonynak tűnhetnek, de tekintve, hogy három osztállyal dolgoztunk (EKZ, korai AK és kontroll) és az alanyok eloszlása is kiegyenlített volt, találgatással csak 33,3%-ot érhetünk el. Ennek figyelembe vételével a manuálisan számított jellemzők elfogadható pontossághoz vezettek: a 61,3%-os érték ezt jelentősen meghaladja. A feladatot kétoosztályos detektálási feladatként értelmezve, az előálló pontosságértékek elég magasak, ami azt jelzi, hogy a téves osztályozások nagy része az EKZ és korai AK osztályok keveréséből adódhatott.

Ugyanezen (*alapvető*) jellemzők automatikus meghatározásával is 61,3%-os (háromosztályos) pontosságértéket kaptunk, bár a kétoosztályos detektálási metrikák valamivel alacsonyabbnak adódtak. A kibővített jellemzőkészlet használatával viszont az összes pontosságérték nőtt (a fedés kivételével, mely változatlan maradt): bár a (háromosztályos) pontosság csak 1,4%-kal növekedett, az F -érték 2,4%-os (abszolút) növekedése azt jelzi, hogy ezzel a jellemzőkészlettel nagyobb pontosságú demenciadetektálásra van lehetőség. Ráadásul ezen jellemzőkészlet manuális kiszámítása irreálisan sok erőfeszítést igényelt volna, míg automatikusan csak elhanyagolhatóan nőtt a futási idő.

A 4. táblázat alapján a kibővített jellemzőkészlettel érhetőek el a legjobb eredmények. Ugyanakkor nem minden hiba feltétlenül azonos súlyú; érdemes lehet azt is megvizsgálni, hogy az egyes beszélőcsoportok milyen hatékonysággal

Jellemzőkészlet	Pont.	UAR	Prec.	Fedés	F_1
Kontroll vs. EKZ + eAK	74,7%	72,0%	81,6%	80,0%	80,8%
Kontroll vs. EKZ	76,0%	76,0%	84,2%	64,0%	72,7%
Kontroll vs. eAK	84,0%	84,0%	81,5%	88,0%	84,6%
EKZ vs. eAK	76,0%	76,0%	—	—	—

5. táblázat. A különböző kétosztályos gépi tanulási esetekben elért pontosságértékek, a kibővített akusztikai jellemzőkészlet használata mellett.

különíthetők el egymástól. Habár a tévesztési mátrix elemzésével is vonhatunk le erről következtetéseket, szerintünk ennél jobb módszer az egyes osztálypárokra bináris osztályozó modelleket tanítani és azok teljesítményét vizsgálni. Így utolsó teszt sorozatunkban minden lehetséges osztálypárra kétosztályos SVM modelleket tanítottunk. Az első teszt esetben mind a 75 beszélő adatait felhasználtuk, de az EKZ-val és korai AK-val diagnosztizált pácienseket ugyanabba az osztályba (*demens alanyok*) soroltuk, míg a másik osztály a kontroll csoport volt. A fennmaradó három esetben csak két-két beszélőcsoport 25-25 alanyának felvételeit használtuk. Mivel a korábbi kísérletben a kibővített jellemzőkészlet vezetett a legjobb eredményekhez, most is ezeket a jellemzővektorokat használtuk.

Az 5. táblázat tartalmazza a bináris osztályozó modellek használatával kapott eredményeket. Mikor mind a 75 beszélő adatait használtuk, az eredmények hasonlóan adódtak, mint mikor hámosztályos SVM modelleket tanítottunk és a predikciókban a két demens osztály címkéit összevontuk: a megegyező (74,7%) pontosságérték mellett valamivel alacsonyabb UAR-t és pontosságot (precision) kaptunk, míg a fedés és az F -érték kicsit nőtt. A további eredményeket megvizsgálva azt láthatjuk, hogy a kontroll csoportba tartozó alanyokat volt a legkönnyebb elkülöníteni az enyhe AK-s páciensektől, amely elég logikus, hiszen ez a két csoport található a vizsgált mentális betegségskála két végén. Ehhez képest annak eldöntése, hogy az alany EKZ-ben szenved-e vagy kontroll, illetve hogy EKZ-ben vagy enyhe AK-ban szenved-e, hasonlóan nehéznek bizonyult, bár a 76%-os pontosságértékek véleményünk szerint hatékony elkülönítést tükröznek. (Mikor az EKZ és enyhe AK csoportokba tartozó alanyokat próbáltuk megkülönböztetni egymástól, nem számítottunk ki minden pontosságértéket: mivel ebben az esetben nem volt kézenfekvő pozitív osztályunk, nem láttuk értelmét az információ-visszakeresési jellemzőknek.)

Összességében az eredmények azt tükrözik, hogy a kifejlesztett jellemzőkészlet és a javasolt kétlépéses eljárás sikeresen alkalmazható mind korai enyhe kognitív zavar, mind korai Alzheimer-kór felismerésére. Célunk, hogy az akusztikus jellemzőkre épülő elemzést a későbbiekben megfelelő nyelvészeti jellemzőkkel is kiegészítsük, hiszen a beszélő bizonytalanságát egyes kifejezések (pl. „azt hiszem”, „úgy gondolom”) jelenléte is jelzi (ld. pl. [37]). Ez azonban egyértelműen további kutatásokat igényel.

5. Összegzés

A különböző demenciátípusok, így például az enyhe kognitív zavar és az Alzheimer-kór korai felismerése kritikus a kezelés mielőbbi megkezdése, és így a páciens életminősége szempontjából. Korábbi tanulmányainkban már javasoltunk egy manuális eljárást, mely a beszéd temporális paramétereinek (beszédtempó, artikulációs tempó, hezitációs ráta stb.) vizsgálatával kísérli meg elkülöníteni az EKZ-ban szenvedő alanyokat az egészséges kontrolloktól, majd javasoltunk egy beszédfelismerésen alapuló eljárást az akusztikus jellemzők automatikus kiszámítására is. A diagnózis-hipotézist második lépésben, ezen jellemzők alapján valamely gépi tanulási eljárás szolgáltatta. Jelen dolgozatunkban azt vizsgáltuk, hogy ezek a technikák alkalmasak-e az EKZ mellett a korai AK felismerésére is. Eredményeink azt tükrözik, hogy az eljárás használatával hatékonyan (80% fölötti F-értékkel) meg lehet határozni, hogy az alany rendelkezik-e bármilyen demenciával. A három beszélőcsoport (EKZ, korai AK és kontroll) páronkénti elkülönítésére is 72 – 84% közötti pontosságértékek adódtak; nem meglepő módon az enyhe AK és a kontroll csoportot volt a legkönnyebb elkülöníteni egymástól, mivel ez a két csoport található a vizsgált mentális betegségskála két végén.

Köszönetnyilvánítás

A kutatást az EFOP-3.6.1- 16-2016-00008 azonosítójú, EU társfinanszírozású projekt támogatta. Tóth László kutatásait az MTA Bolyai János ösztöndíja támogatta. Vincze Veronika kutatásait az Emberi Erőforrások Minisztériuma ÚNKP-17-4 kódszámú Új Nemzeti Kiválóság Programja támogatta.

Hivatkozások

1. Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., Ferri, C.: The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's and Dementia* **9**(1) (2013) 63–75
2. Cummings, L., Mega, S.: *Neuropsychiatry and Behavioral Neuroscience*. Oxford University Press (2003)
3. Reichman, W., Cummings, J.: Dementia. In Duthie, E., Katz, P., eds.: *Practice of Geriatrics*. W.B. Saunders, Philadelphia, PA (1999) 395–304
4. Turi, Zs., Németh, D., Hoffmann, I.: Nyelv és emlékezet. In Pléh, Cs., Lukács, Á., eds.: *Pszicholingvisztika 1-2*. Akadémiai Kiadó, Budapest (2014) 743–776
5. Folstein, M., Folstein, S., McHugh, P.: Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12**(3) (1975) 189–198
6. Freedman, M., Leach, L., Kaplan, E., Winocur, G., Shulman, K., Delis, D.: *Clock Drawing: A Neuropsychological Analysis*. New York: Oxford University Press (1994)
7. Kálmán, J., Maglóczy, E., Janka, Z.: Óra Rajzolási Teszt: gyors és egyszerű dementia szűrő módszer. *Psychiatria Hungarica* **10**(3) (1995) 11–18

8. Rosen, W., Mohs, R., Davis, K.: A new rating scale for Alzheimer's disease. *Journal of Psychiatric Research* **141**(11) (1984) 1356–1364
9. Roark, B., Mitchell, M., Hosom, J.P., Hollingshead, K., Kaye, J.: Spoken language derived measures for detecting mild cognitive impairment. *IEEE Transactions on Audio, Speech, and Language Processing* **19**(7) (2011) 2081–2090
10. Forbes-McKay, K., Venneri, A.: Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *International Journal of Geriatric Psychiatry* **2005**(26) (2005) 243–256
11. Hoffmann, I., Németh, D., Dye, C.D., Pákási, M., Irinyi, T., Kálmán, J.: Temporal parameters of spontaneous speech in Alzheimer's disease. *International Journal of Speech-Language Pathology* **12**(1) (2010) 29–34
12. Meilán, J.J.G., Martínez-Sánchez, F., Carro, J., Sánchez, J.A., Pérez, E.: Acoustic markers associated with impairment in language processing in Alzheimer's disease. *Spanish Journal of Psychology* **15**(2) (2012) 487–494
13. Satt, A., Hoory, R., König, A., Aalten, P., Robert, P.H.: Speech-based automatic and robust detection of very early dementia. In: *Interspeech, Szingapúr* (2014) 2538–2542
14. Jarrold, W., Peintner, B., Wilkins, D., Vergryi, D., Richey, C., Gorno-Tempini, M.L., Ogar, J.: Aided diagnosis of dementia type through computer-based analysis of spontaneous speech. In: *CLPsych, Baltimore, USA* (2014) 27–37
15. Laske, C., Sohrabi, H.R., Frost, S.M., de Ipiña, K.L., Garrard, P., Buscema, M., Dauwels, J., Soekadar, S.R., Mueller, S., Linnemann, C., Bridenbaugh, S.A., Kanagasangam, Y., Martins, R.N., O'Bryant, S.E.: Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer's & Dementia* **11**(5) (2015) 561–578
16. Garrard, P., Maloney, L.M., Hodges, J.R., Patterson, K.: The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain* **128**(2) (2005) 250–260
17. Taler, V., Phillips, N.: Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology* **30**(5) (2008) 501–556
18. Santos, V.D., Thomann, P.A., Wüstenberg, T., Seidl, U., Essig, M., Schröder, J.: Morphological cerebral correlates of CERAD test performance in mild cognitive impairment and Alzheimer's disease. *Journal of Alzheimer's Disease* **23**(3) (2011) 411–420
19. Cardoso, S., Silva, D., Maroco, J., de Mendonca, A., Guerreiro, M.: Non-literal language deficits in Mild Cognitive Impairment. *Psychogeriatrics* **14**(4) (2014) 222–228
20. Garrard, P., Rentoumi, V., Gesierich, B., Miller, B., Gorno-Tempini, M.L.: Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex* **55** (2014) 122–129
21. Barth, S., Schönknecht, P., Pantel, J., Schröder, J.: Mild cognitive impairment and Alzheimer's disease: an investigation of the CERAD-NP test battery. *Fortschritte der Neurologie-Psychiatrie* **73**(10) (2005) 568–576
22. Tóth, L., Gosztolya, G., Vincze, V., Hoffmann, I., Szatlóczki, G., Bíró, E., Zsura, F., Pákási, M., Kálmán, J.: Automatic detection of mild cognitive impairment from spontaneous speech using ASR. In: *Interspeech, Drezda, Németország* (2015) 2694–2698
23. Tóth, L., Gosztolya, G., Vincze, V., Hoffmann, I., Szatlóczki, G., Bíró, E., Zsura, F., Pákási, M., Kálmán, J.: Az enyhe kognitív zavar automatikus azonosítása. In: *Neumann Kollokvium, Veszprém, Magyarország* (2015) 112–117

24. Gosztolya, G., Tóth, L., Grósz, T., Vincze, V., Hoffmann, I., Szatlóczki, G., Pákáski, M., Kálmán, J.: Detecting Mild Cognitive Impairment from spontaneous speech by correlation-based phonetic feature selection. In: Interspeech, San Francisco, USA (2016) 107–111
25. Baum, S.R., Blumstein, S.E., Naeser, M.A., Palumbo, C.L.: Temporal dimensions of consonant and vowel production: An acoustic and CT scan analysis of aphasic speech. *Brain and Language* **39**(1) (1990) 33–56
26. Illes, J.: Neurolinguistic features of spontaneous language production dissociate three forms of neurodegenerative disease: Alzheimer's, Huntington's, and Parkinson's. *Brain and language* **37**(4) (1989) 628–642
27. Juncos-Rabadán, O., Pereiro, A.X., Facaly, D., Rodríguez, N.: Una revisión de la investigación sobre lenguaje en el deterioro cognitivo leve. *Revista de Logopedia, Foniatría y Audiología* **30**(2) (2010) 73–83
28. Gósy, M.: The paradox of speech planning and production. *Magyar Nyelvőr* **12**(1) (1998) 3–15
29. Tóth, L.: Phone recognition with hierarchical Convolutional Deep Maxout Networks. *EURASIP Journal on Audio, Speech, and Music Processing* **2015**(25) (2015) 1–13
30. Sachdev, P., Lipnicki, D., Crawford, J., Reppermund, S., Kochan, N., Trollor, J., Draper, B., Slavin, M., Kang, K., Lux, O., Brodaty, K.M.H.: Risk profiles for mild cognitive impairment vary by age and sex: the Sydney memory and ageing study. *The American Journal of Geriatric Psychiatry* **20**(10) (2012) 854–865
31. Kockmann, M., Burget, L., Cernocký, J.: Brno University of Technology system for Interspeech 2010 Paralinguistic Challenge. In: Interspeech, Makuhari, Chiba, Japán (2010) 2822–2825
32. Grzybowska, J., Kacprzak, S.: Speaker age classification and regression using i-vectors. In: Interspeech, San Francisco, USA (2016) 1402–1406
33. Gósy, M.: BEA: A multifunctional Hungarian spoken language database. *The Phonetician* **105**(106) (2012) 50–61
34. Lehr, M., Prud'hommeaux, E., Shafran, I., Roark, B.: Fully automated neuropsychological assessment for detecting Mild Cognitive Impairment. In: Interspeech, Portland, USA (2012) 1039–1042
35. Schölkopf, B., Platt, J., Shawe-Taylor, J., Smola, A., Williamson, R.: Estimating the support of a high-dimensional distribution. *Neural Computation* **13**(7) (2001) 1443–1471
36. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter* **11**(1) (2009) 10–18
37. Vincze, V., Hoffmann, I., Szatlóczki, G., Bíró, E., Gosztolya, G., Tóth, L., Pákáski, M., Kálmán, J.: Az enyhe kognitív zavar automatikus azonosítása beszédátiratok alapján. In: MSZNY, Szeged (2015) 249–256