Depression Speaks
Automatic Discrimination Between Depressed and Non-Depressed Speakers Based on Nonverbal Speech Features

F.Scibelli, G.Roffo, M.Tayarani, L.Bartoli, G.De Mattia, A.Esposito and A.Vinciarelli

Università degli Studi della Campania (Italy)
Unità Operative Salute Mentale (Italy)
University of Glasgow (UK)
Outline

• Introduction
• The Data
• The Approach
• Experiments and Results
• Conclusions
Outline

• Introduction
• The Data
• The Approach
• Experiments and Results
• Conclusions
Why Depression?
Why Depression?

- 300 million patients in 2015 (World Health Organisation)
Why Depression?

• 300 million patients in 2015 (World Health Organisation)

• Second cause of disability after ischaemia (European Commission)
Why Depression?

• 300 million patients in 2015 (World Health Organisation)

• Second cause of disability after ischaemia (European Commission)

• Most important suicide factor for elderly people (European Commission)
Why Nonverbal (Speech) Behaviour?
Why Nonverbal (Speech) Behaviour?

- Traditional methods (HRSD and BDI-II) not fully robust to biases of both clinicians and patients
Why Nonverbal (Speech) Behaviour?

• Traditional methods (HRSD and BDI-II) not fully robust to biases of both clinicians and patients

• Detection of nonverbal behavioural markers can limit the effect of biases while reducing the costs
Outline

- Introduction
- The Data
- The Approach
- Experiments and Results
- Conclusions
Depression

Control
Depression

Control

20
Depression

20 42

Control
Depression

| 20 | 42 |

Control

| 20 |    |
Depression

20  42

Control

20  37
DSM-based diagnosis by professional clinicians
Depression

DSM-based diagnosis by professional clinicians

20 42

Control

Match based on gender, age and education

20 37
DSM-based diagnosis by professional clinicians

Match based on gender, age and education

BDI-II Questionnaire
Depression

20
42

DSM-based diagnosis by professional clinicians

BDI-II Questionnaire

62 Diary recordings
57 Tale recordings

Control

20
37

Match based on gender, age and education
Depression

- DSM-based diagnosis by professional clinicians
- BDI-II Questionnaire
- 62 Diary recordings
- 57 Tale recordings

Control

- Match based on gender, age and education
- 52 Diary recordings
- 54 Tale recordings
Diary is an interview about the activities of the last weekend (spontaneous speech)

Depression

20
42

Control

20
37

62 Diary recordings
57 Tale recordings

52 Diary recordings
54 Tale recordings
Diary is an interview about the activities of the last week end (spontaneous speech)

Depression

20
42

Control

20
37

62 Diary recordings
57 Tale recordings

52 Diary recordings
54 Tale recordings
Diary is an interview about the activities of the last weekend (spontaneous speech)

Depression

20
42

Control

20
37

62 Diary recordings
57 Tale recordings

52 Diary recordings
54 Tale recordings
Tale is reading a short story written by Aesopus (read speech)

Depression

20
42

Control

20
37

62 Diary recordings
57 Tale recordings

52 Diary recordings
54 Tale recordings
Tale is reading a short story written by Aesopus (read speech)

Depression
20 42

Control
20 37

62 Diary recordings
57 Tale recordings

52 Diary recordings
54 Tale recordings
**Tale** is reading a short story written by Aesop (read speech)

- **Depression**
  - 20 Diary recordings
  - 42 Tale recordings

- **Control**
  - 20 Diary recordings
  - 37 Tale recordings

62 Diary recordings
57 Tale recordings

52 Diary recordings
54 Tale recordings
Outline

- Introduction
- The Data
- The Approach
- Experiments and Results
- Conclusions
Speech Segmentation

Feature Extraction

Feature Selection

Classification
Speech Segmentation

Feature Extraction

Feature Selection

Classification

Selection of patient or control speech
Speech Segmentation

Feature Extraction

Feature Selection

Classification

Selection of patient or control speech

Extraction of 384 features (IS09)

openSMILE by audEERING™
Speech Segmentation

Feature Extraction

Feature Selection

Classification

Selection of patient or control speech

Extraction of 384 features (IS09)

Infinite Latent Feature Selection

openSMILE:}

by audEERING™
Speech Segmentation

Feature Extraction

Feature Selection

Classification

Selection of patient or control speech

Extraction of 384 features (IS09)

Infinite Latent Feature Selection

Identification of markers
Speech Segmentation → Feature Extraction → Feature Selection → Classification

Selection of patient or control speech

Extraction of 384 features (IS09)

Infinite Latent Feature Selection

Identification of markers

Support Vector Machines
Speech Segmentation → Feature Extraction → Feature Selection → Classification

Selection of patient or control speech

Extraction of 384 features (IS09)

Infinite Latent Feature Selection

Identification of markers

Support Vector Machines

Depressed vs Control
Outline

• Introduction
• The Data
• The Approach
• Experiments and Results
• Conclusions
Leave One Subject Out approach for both Linear Kernel SVM and Infinite Latent Feature Selection
<table>
<thead>
<tr>
<th>Task</th>
<th>Precision</th>
<th>Recall</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diary</td>
<td>66%</td>
<td>60%</td>
<td>68%</td>
</tr>
<tr>
<td>Tale</td>
<td>75%</td>
<td>74%</td>
<td>76%</td>
</tr>
<tr>
<td>Diary-FS</td>
<td>74%</td>
<td>65%</td>
<td>74%</td>
</tr>
<tr>
<td>Tale-FS</td>
<td>74%</td>
<td>80%</td>
<td>77%</td>
</tr>
<tr>
<td></td>
<td>Diary</td>
<td>Diary-FS</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>D</td>
<td>74.2%</td>
<td>25.8%</td>
<td>80.62%</td>
</tr>
<tr>
<td>C</td>
<td>40.4%</td>
<td>59.6%</td>
<td>34.6%</td>
</tr>
<tr>
<td></td>
<td>Tale</td>
<td>Tale-FS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>D</td>
<td>77.2%</td>
<td>22.8%</td>
<td>73.7%</td>
</tr>
<tr>
<td>C</td>
<td>25.9%</td>
<td>74.1%</td>
<td>20.4%</td>
</tr>
</tbody>
</table>
Outline

• Introduction
• The Data
• The Approach
• Experiments and Results
• Conclusions
Conclusions
Conclusions

- While being preliminary, the approach discriminates well above chance
Conclusions

• While being preliminary, the approach discriminates well above chance

• Unlike previous works, the approach performs better on read speech (in line with the psychiatric literature)
Conclusions

- While being preliminary, the approach discriminates well above chance
- Unlike previous works, the approach performs better on read speech (in line with the psychiatric literature)
- Future work includes the adoption of Deep Networks and the use of the transcriptions
Thank You!

F.Scibelli, G.Roffo, M.Tayarani, L.Bartoli, G.De Mattia, A.Esposito and A.Vinciarelli

Università’ degli Studi della Campania (Italy)
Unita’ Operative Salute Mentale (Italy)
University of Glasgow (UK)