DRUM TRANSCRIPTION VIA JOINT BEAT AND DRUM MODELING USING CONVOLUTIONAL RNNs

Richard Vogl\textsuperscript{1,2}, Matthias Dorfer\textsuperscript{2}, Gerhard Widmer\textsuperscript{2}, Peter Knees\textsuperscript{1}

richard.vogl@tuwien.ac.at, matthias.dorfer@jku.at, gerhard.widmer@jku.at, peter.knees@tuwien.ac.at
WHAT IS DRUM TRANSCRIPTION?

■ Input: western popular music containing drums
■ Output: symbolic representation of notes played by drum instruments
WHAT IS DRUM TRANSCRIPTION?

Focus on the three major drum instruments:
- bass or kick drum (KD)
- snare drum (SD)
- hi-hat (HH)

Reasons:
- Dominant instruments: most onsets
- Common subset for public datasets
SYSTEM OVERVIEW

Audio signals are preprocessed, followed by feature extraction and event detection. Once events are detected, they are picked for peak picking. The output of peak picking is used for NN training, which in turn leads to the classification of audio events.
SYSTEM OVERVIEW

Audio signal preprocessing → NN feature extraction
event detection → classification → peak picking → events

- spectrogram: f [Hz] vs. t [s]
- NN training
SYSTEM OVERVIEW

- **Signal preprocessing**
- **Feature extraction**
- **Event detection**
- **Classification**
- **NN training**
- **Peak picking**

**Input:** Audio signal

**Output:** Audio events

- **Spectrogram:** Frequency (f [Hz]) vs. Time (t [s])
- **Activation functions:**
  - Hi-hat
  - Snare
  - Kick
SYSTEM OVERVIEW

**Audio**

1. Signal preprocessing
2. Feature extraction
3. Event detection
4. Classification
5. Peak picking
6. NN training

**Spectrogram**

- $f \text{ [Hz]}$
- $t \text{ [s]}$

**Events**

- hi-hat
- snare
- kick

**Activation functions**
ISSUES OF CURRENT SYSTEMS
ISSUES OF CURRENT SYSTEMS

- Performance not satisfying on real music
ISSUES OF CURRENT SYSTEMS

- Performance not satisfying on real music
- Do not produce additional information for transcripts
  
  *drum onset detection vs drum transcription*
ISSUES OF CURRENT SYSTEMS

- Performance not satisfying on real music
- Do not produce additional information for transcripts

*drum onset detection* vs *drum transcription*

- bars lines
ISSUES OF CURRENT SYSTEMS

- Performance not satisfying on real music
- Do not produce additional information for transcripts
  *drum onset detection vs drum transcription*
  - bars lines
  - tempo
Performance not satisfying on real music

Do not produce additional information for transcripts

\textit{drum onset detection} vs \textit{drum transcription}

- bars lines
- tempo
- meter
ISSUES OF CURRENT SYSTEMS

- Performance not satisfying on real music
- Do not produce additional information for transcripts

*drum onset detection* vs *drum transcription*

- bars lines
- tempo
- meter
- dynamics / accents
ISSUES OF CURRENT SYSTEMS

- Performance not satisfying on real music
- Do not produce additional information for transcripts
  *drum onset detection vs drum transcription*
  - bars lines
  - tempo
  - meter
  - dynamics / accents
  - stroke / playing technique
ISSUES OF CURRENT SYSTEMS

- Performance not satisfying on real music
- Do not produce additional information for transcripts
  *drum onset detection vs drum transcription*
    - bars lines
    - tempo
    - meter
    - dynamics / accents
    - stroke / playing technique
- Only three instrument classes
- etc.
ISSUES OF CURRENT SYSTEMS

- Performance not satisfying on real music
- Do not produce additional information for transcripts
  - drum onset detection vs drum transcription
    - bars lines
    - tempo
    - meter
    - dynamics / accents
    - stroke / playing technique
- Only three instrument classes
- etc.
ADDITIONAL INFORMATION FOR TRANSCRIPTS
Use beat and downbeat tracking to get:
Use **beat and downbeat tracking** to get:

- bars lines
- tempo
- meter
Use **beat and downbeat tracking** to get:

- bars lines
- tempo
- meter
IMPROVE PERFORMANCE

Three components to reach this goal:
1. Leverage beat information
2. Better model for drum detection
3. Dataset with real music for training
1. LEVERAGE BEAT INFORMATION
1. LEVERAGE BEAT INFORMATION

Beats are **highly correlated** with drum patterns.
1. LEVERAGE BEAT INFORMATION

Beats are **highly correlated** with drum patterns.

Assume that **prior knowledge** of beats is helpful for drum transcription (drum hit locations / repetitive patterns)
1. LEVERAGE BEAT INFORMATION

- Beats are **highly correlated** with drum patterns
- Assume that **prior knowledge** of beats is helpful for drum transcription (drum hit locations / repetitive patterns)
- Use **multi-task learning** for beats and drums
MULTI-TASK LEARNING

input

output

\[ f \text{ [Hz]} \]

\[ t \text{ [s]} \]
MULTI-TASK LEARNING

Three experiments:

- input

- output
MULTI-TASK LEARNING

Three experiments:
- Training on drum targets ($DT$)
MULTI-TASK LEARNING

Three experiments:

- Training on drum targets ($DT$)
- Training on drum targets with annotated beats as additional input features ($BF$)
MULTI-TASK LEARNING

Three experiments:

- Training on drum targets (DT)
- Training on drum targets with annotated beats as additional input features (BF)
- Training on drum and beat targets as multi-task problem (MT)
MULTI-TASK LEARNING

Three experiments:
- Training on drum targets (DT)
- Training on drum targets with annotated beats as additional input features (BF)
- Training on drum and beat targets as multi-task problem (MT)

Expected increase in performance for BF compared to DT
MULTI-TASK LEARNING

Three experiments:

- Training on drum targets \((DT)\)
- Training on drum targets with annotated beats as additional input features \((BF)\)
- Training on drum and beat targets as multi-task problem \((MT)\)

Expected increase in performance for \(BF\) compared to \(DT\)

Expected increase in performance for \(MT\) compared to \(DT\)
2. NETWORK MODELS — BASELINE MODELS
2. NETWORK MODELS — BASELINE MODELS

- Recurrent neural networks
2. NETWORK MODELS — BASELINE MODELS

- Recurrent neural networks
  - Recurrent connections act as **memory**
  - Processing of **sequential data**

RNN train data sample
2. NETWORK MODELS — BASELINE MODELS

- Recurrent neural networks
  - Recurrent connections act as memory
  - Processing of sequential data
  - Work well for drum detection and beat tracking
    [Böck et al. ISMIR’16]
2. NETWORK MODELS — BASELINE MODELS

- Recurrent neural networks
  - Recurrent connections act as memory
  - Processing of sequential data
  - Work well for drum detection and beat tracking [Böck et al. ISMIR’16]

- RNN with label time shift (tsRNN)
  - state-of-the-art baseline [Vogl et al. ICASSP’17]

- Bidirectional recurrent NN (BDRNN)
  - [Vogl et al. ISMIR’16] [Southall et al. ISMIR’16]
    - Similar performance tsRNN
2. NETWORK MODELS — NEW FOR DT
2. NETWORK MODELS — NEW FOR DT

- Convolutional NN (CNN)
  - Convolutions capture **local correlations**
  - **Acoustic modeling** of drum sounds

![CNN train data sample](image)
2. NETWORK MODELS — NEW FOR DT

- Convolutional NN (CNN)
  - Convolutions capture **local correlations**
  - **Acoustic modeling** of drum sounds

- Convolutional BDRNN (CRNN)
  - "best of both worlds"
  - Low-level CNN for **acoustic modeling**
  - Higher-level RNN for **repetitive pattern modeling**
## NETWORK MODELS

<table>
<thead>
<tr>
<th></th>
<th>Frames</th>
<th>Context</th>
<th>Conv. Layers</th>
<th>Rec. Layers</th>
<th>Dense Layers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BDRNN (S)</strong></td>
<td>100</td>
<td>—</td>
<td>—</td>
<td>2x50 GRU</td>
<td>—</td>
</tr>
<tr>
<td><strong>BDRNN (L)</strong></td>
<td>400</td>
<td>—</td>
<td>—</td>
<td>3x30 GRU</td>
<td>—</td>
</tr>
<tr>
<td><strong>CNN (S)</strong></td>
<td>—</td>
<td>9</td>
<td>2 x 32 3x3 filt. 3x3 max pooling 2 x 64 3x3 filt.</td>
<td>—</td>
<td>2x256</td>
</tr>
<tr>
<td><strong>CNN (L)</strong></td>
<td>—</td>
<td>25</td>
<td>3x3 max pooling all w/ batch norm.</td>
<td>—</td>
<td>2x256</td>
</tr>
<tr>
<td><strong>CRNN (S)</strong></td>
<td>100</td>
<td>9</td>
<td>2x50 GRU</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td><strong>CRNN (L)</strong></td>
<td>400</td>
<td>13</td>
<td>3x60 GRU</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*tsRNN* state-of-the-art baseline [Vogl et al. ICASSP’17]
CLASSIC DATASETS (ONLY DRUMS)
CLASSIC DATASETS (ONLY DRUMS)

- IDMT-SMT-Drums [Dittmar and Gärtner 2014]
  - Solo drum tracks, recorded, synthesized, and sampled
  - 95 tracks, total: **24m**, onsets: **8004** + training samples
CLASSIC DATASETS (ONLY DRUMS)

- **IDMT-SMT-Drums** [Dittmar and Gärtner 2014]
  - Solo drum tracks, recorded, synthesized, and sampled
  - 95 tracks, total: **24m**, onsets: **8004** + training samples
CLASSIC DATASETS (ONLY DRUMS)

- **IDMT-SMT-Drums** [Dittmar and Gärtner 2014]
  - Solo drum tracks, recorded, synthesized, and sampled
  - 95 tracks, total: 24m, onsets: 8004 + training samples

- **ENST-Drums** [Gillet and Richard 2006]
  - Recordings, three drummers on different drum kits, optional accompaniment
  - 64 tracks, total: 1h, onsets: 22391 + training samples
CLASSIC DATASETS (ONLY DRUMS)

- **IDMT-SMT-Drums** [Dittmar and Gärtner 2014]
  - Solo drum tracks, recorded, synthesized, and sampled
  - 95 tracks, total: **24m**, onsets: **8004** + training samples

- **ENST-Drums** [Gillet and Richard 2006]
  - Recordings, three drummers on different drum kits, **optional accompaniment**
  - 64 tracks, total: **1h**, onsets: **22391** + training samples
CLASSIC DATASETS (ONLY DRUMS)

- **IDMT-SMT-Drums** [Dittmar and Gärtner 2014]
  - Solo drum tracks, recorded, synthesized, and sampled
  - 95 tracks, total: 24m, onsets: 8004 + training samples

- **ENST-Drums** [Gillet and Richard 2006]
  - Recordings, three drummers on different drum kits, **optional accompaniment**
  - 64 tracks, total: 1h, onsets: 22391 + training samples
DT 3-FOLD CV RESULTS ON CLASSIC DATASETS
3. NEW DATASETS (DRUMS AND BEATS)

RBMA13-Drums [http://ifs.tuwien.ac.at/~vogl/datasets/]

- Free music from the 2013 Red Bull Music Academy, different styles
- 27 tracks, total: 1h 43m, onsets: 24365
- drum, beat, and downbeat annotations
3. NEW DATASETS (DRUMS AND BEATS)

RBMA13-Drums [http://ifs.tuwien.ac.at/~vogl/datasets/]

- Free music from the 2013 Red Bull Music Academy, different styles
- 27 tracks, total: 1h 43m, onsets: 24365
- drum, beat, and downbeat annotations
3. NEW DATASETS (DRUMS AND BEATS)

**RBMA13-Drums** [http://ifs.tuwien.ac.at/~vogl/datasets/]

- Free music from the 2013 Red Bull Music Academy, different styles
- 27 tracks, total: **1h 43m**, onsets: **24365**
- **drum, beat, and downbeat** annotations
3. NEW DATASETS (DRUMS AND BEATS)

RBMA13-Drums [http://ifs.tuwien.ac.at/~vogl/datasets/]

- Free music from the 2013 Red Bull Music Academy, different styles
- 27 tracks, total: 1h 43m, onsets: 24365
- drum, beat, and downbeat annotations

Multi-task evaluation

- DT: Drum transcription / three fold cross-validation (same as on SMT and ENST)
- BF: Drum transcription using annotated beats as additional input features
- MT: Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13

![Bar chart showing F-measure values for different models and techniques.](chart.png)

- **BDRNN (S)**
- **BDRNN (L)**
- **CNN (S)**
- **CNN (L)**
- **CRNN (S)**
- **CRNN (L)**

- **DT** … Drum transcription (3-fold CV)
- **BF** … Drum transcription using annotated beats as additional input features
- **MT** … Drum transcription and beat detection via multi-task learning

F-measure [%]

- 50
- 55
- 60
- 65
- 70

<table>
<thead>
<tr>
<th>F-measure [%]</th>
<th>BDRNN (S)</th>
<th>BDRNN (L)</th>
<th>CNN (S)</th>
<th>CNN (L)</th>
<th>CRNN (S)</th>
<th>CRNN (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>65</td>
<td>65</td>
<td>60</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>50</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
</tr>
</tbody>
</table>
RESULTS ON RBMA13: BDRNNs

- **F-measure [%]**
  - 50
  - 55
  - 60
  - 65
  - 70

- **BDRNN (S)**
- **BDRNN (L)**

**Legend:**
- **DT** … Drum transcription (3-fold CV)
- **BF** … Drum transcription using annotated beats as additional input features
- **MT** … Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: BDRNNs

Impact on bi-directional RNNs:

- DT ... Drum transcription (3-fold CV)
- BF ... Drum transcription using annotated beats as additional input features
- MT ... Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: BDRNNs

Impact on bi-directional RNNs:
- BF improves for both models ✔

<table>
<thead>
<tr>
<th></th>
<th>DT … Drum transcription (3-fold CV)</th>
<th>BF … Drum transcription using annotated beats as additional input features</th>
<th>MT … Drum transcription and beat detection via multi-task learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDRNN (S)</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
</tr>
<tr>
<td>BDRNN (L)</td>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
</tr>
</tbody>
</table>
RESULTS ON RBMA13: BDRNNs

Impact on bi-directional RNNs:
- BF improves for both models ✓
- MT improves for both models ✓
RESULTS ON RBMA13: BDRNNs

Impact on bi-directional RNNs:
- BF improves for both models ✔
- MT improves for both models ✔
- MT even better than BF for small model !

Impact on bi-directional RNNs:
- DT … Drum transcription (3-fold CV)
- BF … Drum transcription using annotated beats as additional input features
- MT … Drum transcription and beat detection via multi-task learning

![Graph showing F-measure results for BDRNN (S) and BDRNN (L)]
RESULTS ON RBMA13: CNNs

<table>
<thead>
<tr>
<th>Method</th>
<th>F-measure [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT (3-fold CV)</td>
<td>65</td>
</tr>
<tr>
<td>BF</td>
<td>65</td>
</tr>
<tr>
<td>MT</td>
<td>60</td>
</tr>
</tbody>
</table>

- **DT**: Drum transcription (3-fold CV)
- **BF**: Drum transcription using annotated beats as additional input features
- **MT**: Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: CNNs

Impact on CNNs:

- DT … Drum transcription (3-fold CV)
- BF … Drum transcription using annotated beats as additional input features
- MT … Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: CNNs

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent

Impact on CNNs:
- BF inconsistent
RESULTS ON RBMA13: CNNs

Impact on CNNs:
- BF inconsistent
- MT declines for both models
RESULTS ON RBMA13: CRNNs

- DT … Drum transcription (3-fold CV)
- BF … Drum transcription using annotated beats as additional input features
- MT … Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: CRNNs

Impact on CRNNs:

- DT ... Drum transcription (3-fold CV)
- BF ... Drum transcription using annotated beats as additional input features
- MT ... Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: CRNNs

Impact on CRNNs:
- BF improves for both models ✔

Legend:
- DT ... Drum transcription (3-fold CV)
- BF ... Drum transcription using annotated beats as additional input features
- MT ... Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: CRNNs

Impact on CRNNs:
- BF improves for both models ✔
- MT improves for small models ✔

Impact on CRNNs:
- DT … Drum transcription (3-fold CV)
- BF … Drum transcription using annotated beats as additional input features
- MT … Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: CRNNs

Impact on CRNNs:
- BF improves for both models
- MT improves for small models
- MT even better than BF for small model

Impact on CRNNs:

- BF improves for both models
- MT improves for small models
- MT even better than BF for small model

<table>
<thead>
<tr>
<th>Method</th>
<th>CRNN (S)</th>
<th>CRNN (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT (3-fold CV)</td>
<td>65</td>
<td>68</td>
</tr>
<tr>
<td>BF</td>
<td>67</td>
<td>70</td>
</tr>
<tr>
<td>MT</td>
<td>68</td>
<td>72</td>
</tr>
</tbody>
</table>

Legend:
- DT ... Drum transcription (3-fold CV)
- BF ... Drum transcription using annotated beats as additional input features
- MT ... Drum transcription and beat detection via multi-task learning
RESULTS ON RBMA13: CRNNs

Impact on CRNNs:
- BF improves for both models ✓
- MT improves for small models ✓
- MT even better than BF for small model !
- MT equal for large model ?

![Graph showing F-measure (%) for CRNN models]

- DT … Drum transcription (3-fold CV)
- BF … Drum transcription using annotated beats as additional input features
- MT … Drum transcription and beat detection via multi-task learning
RESULTS FOR RECURRENT ARCHITECTURES

<table>
<thead>
<tr>
<th>Architecture</th>
<th>DT</th>
<th>BF</th>
<th>MT</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDRNN (S)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDRNN (L)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRNN (S)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRNN (L)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- DT ... Drum transcription (3-fold CV)
- BF ... Drum transcription using annotated beats as additional input features
- MT ... Drum transcription and beat detection via multi-task learning
RESULTS FOR RECURRENT ARCHITECTURES

F-measure [%]

BDRNN (S)  |  BDRNN (L)  |  CRNN (S)  |  CRNN (L)

DT … Drum transcription (3-fold CV)
BF … Drum transcription using annotated beats as additional input features
MT … Drum transcription and beat detection via multi-task learning
RESULTS FOR RECURRENT ARCHITECTURES

![Graph showing F-measure for different architectures.

- **BDRNN (S)**
- **BDRNN (L)**
- **CRNN (S)**
- **CRNN (L)**

- **DT** ... Drum transcription (3-fold CV)
- **BF** ... Drum transcription using annotated beats as additional input features
- **MT** ... Drum transcription and beat detection via multi-task learning
RESULTS FOR RECURRENT ARCHITECTURES

No improvement because of beat tracking results?

- **BDRNN (S)**
- **BDRNN (L)**
- **CRNN (S)**
- **CRNN (L)**

**Legend:**
- **DT** … Drum transcription (3-fold CV)
- **BF** … Drum transcription using annotated beats as additional input features
- **MT** … Drum transcription and beat detection via multi-task learning
CONCLUSIONS
CONCLUSIONS

- Use beats and downbeats to get **meta information** for transcripts
CONCLUSIONS

- Use beats and downbeats to get meta information for transcripts

- Multi-task learning for drums and beats can be beneficial for recurrent architectures
CONCLUSIONS

- Use beats and downbeats to get **meta information** for transcripts
- **Multi-task learning** for drums and beats can be beneficial for recurrent architectures
- **CRNNs** can outperform RNNs
CONCLUSIONS

- Use beats and downbeats to get **meta information** for transcripts

- **Multi-task learning** for drums and beats can be beneficial for recurrent architectures

- **CRNNs** can outperform RNNs

- **CRNN best overall results** @ MIREX’17 drum transcription
  
  MIREX system: [http://ifs.tuwien.ac.at/~vogl/models/mirex-17.zip](http://ifs.tuwien.ac.at/~vogl/models/mirex-17.zip)
  madmom: [https://github.com/CPJKU/madmom](https://github.com/CPJKU/madmom)
CONCLUSIONS

- Use beats and downbeats to get **meta information** for transcripts

- **Multi-task learning** for drums and beats can be beneficial for recurrent architectures

- **CRNNs** can outperform RNNs

- **CRNN best overall results** @ MIREX’17 drum transcription
  
  MIREX system: [http://ifs.tuwien.ac.at/~vogl/models/mirex-17.zip](http://ifs.tuwien.ac.at/~vogl/models/mirex-17.zip)
  madmom: [https://github.com/CPJKU/madmom](https://github.com/CPJKU/madmom)

- **New dataset** with free music featuring **beat**, and **drum annotations**
  [http://ifs.tuwien.ac.at/~vogl/datasets/](http://ifs.tuwien.ac.at/~vogl/datasets/)
CONCLUSIONS

- Use beats and downbeats to get **meta information** for transcripts

- **Multi-task learning** for drums and beats can be beneficial for recurrent architectures

- **CRNNs** can outperform RNNs

- CRNN **best overall results** @ MIREX’17 drum transcription
  
  MIREX system: [http://ifs.tuwien.ac.at/~vogl/models/mirex-17.zip](http://ifs.tuwien.ac.at/~vogl/models/mirex-17.zip)
  
  madmom: [https://github.com/CPJKU/madmom](https://github.com/CPJKU/madmom)

- **New dataset** with free music featuring **beat**, and **drum annotations**
  
  [http://ifs.tuwien.ac.at/~vogl/datasets/](http://ifs.tuwien.ac.at/~vogl/datasets/)