quEEGNet: Quantum AI for Biosignal Processing

Toshiaki Koike-Akino
Ye Wang

September 28, 2022

MITSUBISHI ELECTRIC RESEARCH LABORATORIES (MERL)
Cambridge, Massachusetts, USA
http://www.merl.com
Outline

• Human-Machine Interaction (HMI)
  – Biosignal processing
  – Brain computer interface (BCI)

• Artificial Intelligence (AI)
  – Deep Learning: Deep Neural Networks (DNN)
  – Post Deep Learning: Quantum Machine Learning (QML)

• Proposal and Validation
  – quEEGNet: Hybrid QNN+DNN solutions
  – Physiological data
  – Experimental validation

• Summary
Brain Machine/Computer Interfaces (BMI/BCI)

- BMI/BCI for reading human’s mind, intention and feelings
- Active researches all over the world
  - Biosignal sensors (ECG, EMG, EEG, ECoG)
  - Robotics, actuators
  - 6G communications
  - Deep learning, Artificial Intelligence (AI)

Artificial Intelligence (AI)

- K-means
- Gaussian mixture model (GMM)
- Principal component analysis (PCA)
- Independent component analysis (ICA)
- Logistic regression (LR)
- Support vector machine (SVM)
- Self-organizing map (SOM)
- Hidden Markov model (HMM)
- Artificial neural networks (ANN)
- Deep learning (DL)
DNN Architectures

- Residual Multi-Layer Perceptron (ResMLP)
- Residual Convolutional Neural Network (ResCNN)
- Bidirectional Long Short-Term Memory (BiLSTM)
- Transformer, U-net, Graph convolution
- Implicit ODE, DEQ, CVX, ...
EEGNet


• Compact CNN yet performing well
Our Publications: Deep Learning for Biosignal Processing


Subject Transfer Learning: Pre-Shot Adversarial Censoring

• Adversarial autoencoder for BCI [Ozdenizci et al, NER’19]
• Complementary adversary [Han et al, SPL’20]
• Rateless autoencoder for soft disentangling [Han et al, JBHI’21]
• Automated Bayesian inference: AutoBayes [Demir et al, Access’21]
• Automated disentanglement: AutoTransfer [Smedemark-Margulies et al, EMBC’22]
  – Rank 1 in subject transfer task of NeurIPS’21 challenge
Moore’s Law: Deep Learning for BMI/BCI

- Number of articles grows exponentially: **doubling every year**
  - More than 10,000 articles

![Graph showing exponential growth of articles with Deep Learning and Deep Neural Net](image_url)

Year


Number of Articles

- 10000
- 1000
- 100
- 10

Deep Learning + EEG
Deep Neural Net + EEG

DL
DNN

Doubling Annually

Red AI: Aiming Higher Performance with Higher-Power Computation

  10.1145/3381831: https://cacm.acm.org/magazines/2020/12/248800-green-ai/

The amount of compute used to train deep learning models has increased 300,000x in six years

Al papers tend to target accuracy rather than efficiency.

© MERL

Deep Learning Crisis for Sustainable Growth

- Escalating power consumption of DNN training
  - [Strubell et al. Energy and policy considerations for deep learning in NLP. 2019]
  - DNN training with network architecture search (NAS) on GPUs requires 5-fold higher carbon emission of single car lifetime!
- New computing modality alternative to CPU/GPU/TPU is desired
  - Natural computing: Quantum computing, DNA computing, etc.

<table>
<thead>
<tr>
<th>Consumption</th>
<th>CO₂e (lbs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air travel, 1 passenger, NY↔SF</td>
<td>1984</td>
</tr>
<tr>
<td>Human life, avg, 1 year</td>
<td>11,023</td>
</tr>
<tr>
<td>American life, avg, 1 year</td>
<td>36,156</td>
</tr>
<tr>
<td>Car, avg incl. fuel, 1 lifetime</td>
<td>126,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Training one model (GPU)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NLP pipeline (parsing, SRL)</td>
<td>39</td>
</tr>
<tr>
<td>w/ tuning &amp; experimentation</td>
<td>78,468</td>
</tr>
<tr>
<td>Transformer (big)</td>
<td>192</td>
</tr>
<tr>
<td>w/ neural architecture search</td>
<td>626,155</td>
</tr>
</tbody>
</table>

Table 1: Estimated CO₂ emissions from training common NLP models, compared to familiar consumption.¹

<table>
<thead>
<tr>
<th>Model</th>
<th>Hardware</th>
<th>Power (W)</th>
<th>Hours</th>
<th>kWh-PUE</th>
<th>CO₂e</th>
<th>Cloud compute cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformer&lt;sub&gt;base&lt;/sub&gt;</td>
<td>P100x8</td>
<td>1415.78</td>
<td>12</td>
<td>27</td>
<td>26</td>
<td>$41–$140</td>
</tr>
<tr>
<td>Transformer&lt;sub&gt;big&lt;/sub&gt;</td>
<td>P100x8</td>
<td>1515.43</td>
<td>84</td>
<td>201</td>
<td>192</td>
<td>$289–$981</td>
</tr>
<tr>
<td>ELMo</td>
<td>P100x3</td>
<td>517.66</td>
<td>336</td>
<td>275</td>
<td>262</td>
<td>$433–$1472</td>
</tr>
<tr>
<td>BERT&lt;sub&gt;base&lt;/sub&gt;</td>
<td>V100x64</td>
<td>12,041.51</td>
<td>79</td>
<td>1507</td>
<td>1438</td>
<td>$3751–$12,571</td>
</tr>
<tr>
<td>BERT&lt;sub&gt;base&lt;/sub&gt;</td>
<td>TPUv2x16</td>
<td>—</td>
<td>96</td>
<td>—</td>
<td>—</td>
<td>$2074–$6912</td>
</tr>
<tr>
<td>NAS</td>
<td>P100x8</td>
<td>1515.43</td>
<td>274,120</td>
<td>656,347</td>
<td>626,155</td>
<td>$942,973–$3,201,722</td>
</tr>
<tr>
<td>NAS</td>
<td>TPUv2x1</td>
<td>—</td>
<td>32,623</td>
<td>—</td>
<td>—</td>
<td>$44,055–$146,848</td>
</tr>
<tr>
<td>GPT-2</td>
<td>TPUv3x32</td>
<td>—</td>
<td>168</td>
<td>—</td>
<td>—</td>
<td>$12,902–$43,008</td>
</tr>
</tbody>
</table>

Table 3: Estimated cost of training a model in terms of CO₂ emissions (lbs) and cloud compute cost (USD).⁷ Power and carbon footprint are omitted for TPUs due to lack of public information on power draw for this hardware.
Quantum Computing

• Morgan Stanley: Quantum tech. can drive 4th industrial revolution
• Quantum processing units (QPU) vendors: IBM, Google, Microsoft, Honeywell, Intel, Nokia, AirBus, IONQ, rigetti, ...
• Quantum cloud services: IBMQ, Amazon Bracket, Microsoft Azure, ...
• Free libraries to evaluate quantum computing on realistic simulators or real devices.
QPU development has been advancing rapidly to allow many qubits

- IBM released **127-qubit** QPUs in Nov. 2021
- IBM plans to release **1121-qubit** QPUs by 2023
Quantum Supremacy

• Some reports claiming to have achieved \textit{quantum supremacy}:
    \begin{itemize}
      \item 53-qubit QPU: $10^9$ faster (200 sec. for 10,000-year job) than classical computers
    \end{itemize}
    \begin{itemize}
      \item Boson sampling: $10^{14}$ faster than classic computer
    \end{itemize}
    \begin{itemize}
      \item Boson sampling: $10^{16}$ faster (36 usec for 9,000-year job) than classical computers
    \end{itemize}

• Quantum advantage is still argued for general applications
Quantum Machine Learning (QML)

- VQE: Variational Quantum Eigensolver
- QAOA: Quantum Approximate Optimization Algorithm
- QAE: Quantum AutoEncoder
- QKSVM: Quantum Kernel Support Vector Machine
- Q(W)GAN: Quantum (Wasserstein) Generative Adversarial Network
- QCNN: Quantum Convolutional Neural Network
- QGNN: Quantum Graph Neural Net
- QX: Quantum Experience
- QCL: Quantum Circuit Learning
- QKS: Quantum Kitchen Sink

Noisy intermediate-scale quantum (NISQ) algorithm
Universal Approximation Theorem/Property (UAT/UAP)

- **UAP for classical neural networks:**
  - Single hidden neural networks can approximate arbitrary bounded continuous functions [Cybenko 1989]
  - Deep hidden neural networks can asymptotically approximate arbitrary functions [Zhou 2017]

- **UAP still holds for quantum computing [Perez 2019]**
  - *Data re-uploading* trick

Wider/deeper neurons can reduce approximation error

\[
    h(\bar{x}) = \sum_{i=1}^{N} \alpha_i \varphi(\bar{w}_i \cdot \bar{x} + b_i)
\]

\[
    U(\bar{x}) = U_N(\bar{x})U_{N-1}(\bar{x}) \cdots U_1(\bar{x}) = \prod_{i=1}^{N} e^{i\omega_i(\phi_i(\bar{x})) \cdot \bar{x}}
\]

![Diagram](image_url)

(a) Neural network  (b) Quantum classifier

Wider neuron → More quantum layers
Deeper layer → More qubits
Differential Programming

• Quantum operation is differentiable:
  – **Parameter shift rule** [Mitarai/Schuld 2018] (exact gradient)
  – Backpropagation through hybrid classical/quantum chips
    – Able to integrate *(implicit) quantum layers* into DNN models
    – e.g., Quanvolutional Neural Network [Henderson2019]
Quantum Neural Network (QNN)

• QML: Post-deep learning paradigm
  – QML is a key major driver for 6G applications [Nawaz et al. Access 2019]

• (Hyped) expectation of QNN advantage:
  – Fewer trainable parameters to support exponentially large quantum states in parallel
  – Parallel ensemble to prevent overfitting and underfitting
  – Low-power processing
Quantum Machine Learning (QML): Moore’s Law

- Number of articles on QML is doubling annually, just **6 years** behind of DNN.
Our Publications: Applied Quantum AI

- Variational quantum algorithm for channel decoding [ISIT19]
- Variational quantum algorithm for demodulation [OFC20]
- Quantum neural network (QNN) to WiFi indoor monitoring [ICC22]
- AutoML to optimize ansatz of QNN [SAM22]
- Variational quantum circuit (VQC) to denoising [ICC22]
- Quantum feature extraction for THz imaging [IRMMW22]
Quantum neural network (QNN) is used to support DNN model (such as EEGNet)
- We call QNN+DNN for biosignal processing as quEEGNet framework for convention
Quantum Neural Network (QNN) for Feature Extraction

• Simplified two-design (S2D) ansatz:
  – Staggered Pauli-Y rotations with controlled Z gates
  – Holding statistical properties identical to ensemble random unitaries with respect to the Haar measure up to the first 2 moments: $SO(2^N) \rightarrow 2N$

• Exponentially large states:
  – 10-qubits to embed 1024-sample signal
Physiological Datasets for Validation

- Stress: temperature, **heart rate**, electrodermal activity, arterial oxygen level, etc. for 4-state stress level measurement
- RSVP: EEG for rapid serial visual presentation (RSVP) drowsiness test with 4 tasks
- MI: PhysioNet EEG Motor Imagery (MI) dataset with 4-class tasks
- ErrP: An error-related potential (ErrP) of EEG dataset in spelling task
- Faces: An implanted electrocorticography (ECoG) array dataset for visual stimulus.
- ASL: An electromyogram (EMG) dataset for fingers motion detection for hand signs.

Heart rates

Electrodermal

Oxygen

RSVP EEG

MI EEG

ErrP EEG

ECoG

EMG

Variety of Datasets

- Publicly available datasets
  - Stress: https://physionet.org/content/noneeg/1.0.0/
  - RSVP: http://hdl.handle.net/2047/D20294523
  - MI: https://physionet.org/physiobank/database/eegmmidb/
  - ErrP: https://www.kaggle.com/c/inria-bci-challenge
  - Faces: https://exhibits.stanford.edu/data/catalog/zk881ps0522
  - ASL: http://hdl.handle.net/2047/D20294523

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Modality</th>
<th>Dimension</th>
<th>Subjects</th>
<th>Classes</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress [48]</td>
<td>Temp. etc.</td>
<td>$7 \times 1$</td>
<td>20</td>
<td>4</td>
<td>24,000</td>
</tr>
<tr>
<td>RSVP [49]</td>
<td>EEG</td>
<td>$16 \times 128$</td>
<td>10</td>
<td>4</td>
<td>41,400</td>
</tr>
<tr>
<td>MI [50]</td>
<td>EEG</td>
<td>$64 \times 480$</td>
<td>106</td>
<td>4</td>
<td>9,540</td>
</tr>
<tr>
<td>ErrP [51]</td>
<td>EEG</td>
<td>$56 \times 250$</td>
<td>27</td>
<td>2</td>
<td>9,180</td>
</tr>
<tr>
<td>Faces Basic [52]</td>
<td>ECoG</td>
<td>$31 \times 400$</td>
<td>14</td>
<td>2</td>
<td>4,100</td>
</tr>
<tr>
<td>Faces Noisy [53]</td>
<td>ECoG</td>
<td>$39 \times 400$</td>
<td>7</td>
<td>2</td>
<td>2,100</td>
</tr>
<tr>
<td>ASL [54]</td>
<td>EMG</td>
<td>$16 \times 50$</td>
<td>5</td>
<td>33</td>
<td>9,900</td>
</tr>
</tbody>
</table>
Performance Results

- quEEGNet achieves state-of-the-art performance yet having few parameters
- Performance improvement via quantum feature extraction for all physiological datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>EEGNet</th>
<th>quEEGNet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress</td>
<td>85.87</td>
<td>87.23</td>
</tr>
<tr>
<td>RSVP</td>
<td>93.73</td>
<td>95.12</td>
</tr>
<tr>
<td>MI</td>
<td>59.61</td>
<td>60.22</td>
</tr>
<tr>
<td>ErrP</td>
<td>74.36</td>
<td>75.92</td>
</tr>
<tr>
<td>Faces Basic</td>
<td>63.30</td>
<td>64.92</td>
</tr>
<tr>
<td>Faces Noisy</td>
<td>75.94</td>
<td>78.01</td>
</tr>
<tr>
<td>ASL</td>
<td>23.64</td>
<td>25.16</td>
</tr>
</tbody>
</table>
Quantum x Biosignal

- Quantum sensing has been revolutionizing biosensing
  - Superconducting quantum interference devices (SQUID)

- Quantum computing for EEG processing

- Our paper is the very first demonstration that applied quantum AI to BCI
Conclusions

• We showed recent AI trends for biosignal processing
• We overviewed recent advancement on quantum AI (QAI) as post-deep learning paradigm
• We introduced the use of emerging QAI for biosignal processing & BCI
  – Demonstrated the first proof-of-concept study for future quantum-era
  – Showed the feasibility of QAI-assisted biosignal processing
  – Achieved state-of-the-art performance with few-parameter QML
  – Showed gain via hybrid QNN + DNN
• There are many fascinating topics and high potentials for future work
• Questions?
  – Please contact me: koike@merl.com