PROJECT REPORT

MODEL COMPLEXITY VS. ACCURACY

An analysis of EEG Motor Imagery classification performance using BCI Competition IV 2a dataset.

Joseph Hong

PROBLEM STATEMENT

THE PROBLEM

There is an assumption that increased model complexity/depth correlates with an improvement of classification accuracy, under the premise of readily available data.

However, in the context of BCI, particularly **motor imagery-based BCI**, generalization is just as important as personalization.

What are the tradeoffs of model complexity in the context of motor imagery classification?

THE CHALLENGE

EEG motor imagery data poses an interesting challenge as it has:

- ⚡ Low Signal-to-Noise Ratio (SNR)
- Extremely High Inter-Subject variability
- Non-stationary signals
- ✖ Lack of available data (vs images, text, etc.)
- MI for different body parts can be very similar

DATASET

BCI Competition IV 2a

- **Task:** 4-Class Motor Imagery (Left Hand, Right Hand, Feet, Tongue).
- **Subjects:** 9 distinct individuals.
- **Data:** 22 EEG channels, sampled at 250 Hz.
- **Volume:** 288 trials per subject across 2 sessions.

EEG Preprocessing

- **EEG Only:** Removed EOG channels
- **BPF:** 4-38 Hz (skipped notch since it would be >50)
- **Sub-Band (FBCSP):** Additional bands (8–12), (12–16), (16–20), (20–24), (24–28), (28–32)

EVALUATION

Leave-One-Subject-Out

- **Generalization:** How well do the models generalize across subjects?
- **Training Set: 100%** samples from other subjects
- **Validation Set: 100%** samples from **one** subject

Within Subject (K-Fold, K = 3)

- **Personalization:** How well can models discern subtle differences in MI within subjects?
- **Training Set: 80%** samples from a subject
- **Validation Set: 20%** samples from same subject

METHODS: THE MODELS

FBCSP + SVM

Linear

Filter Bank Common Spatial Patterns. Breaking down data into frequency band-level features, then feeding them into a linear Support Vector Machine.

Parameters: 2 CSP filters, 20 best components

EEGNET

Simple DL

Compact Convolutional Neural Network. Uses depthwise separable convolutions to learn spatial and temporal filters efficiently.

Hyperparameters: 60 epochs, 64 batch size, 1e-3 learning rate

ATCNET

Complex DL

Attention Temporal Convolutional Network. Combines high-capacity TCN blocks with multi-head attention mechanisms.

Hyperparameters: 80 epochs, 32 batch size, 5e-4 learning rate

The full hyperparameters can be found in the code.

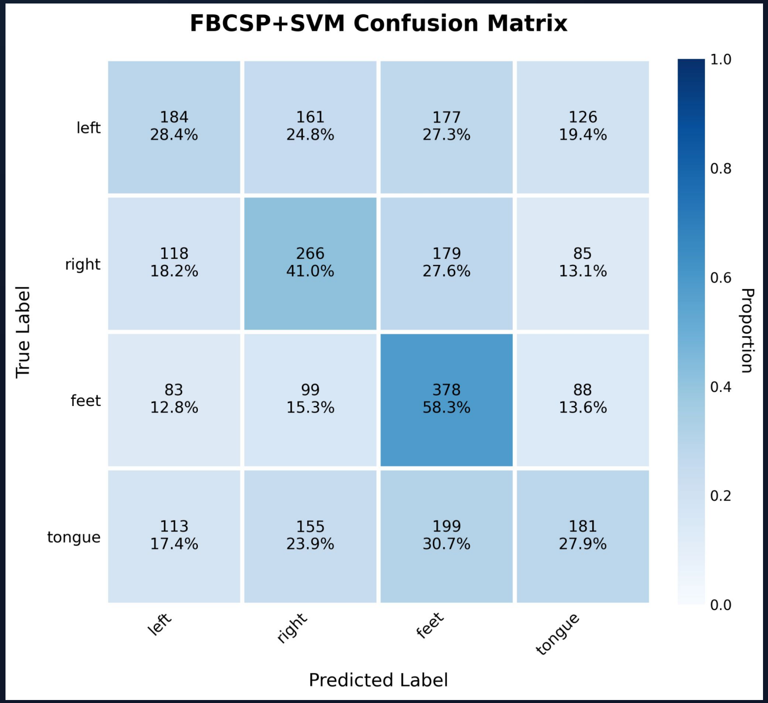
COMPLEXITY ANALYSIS

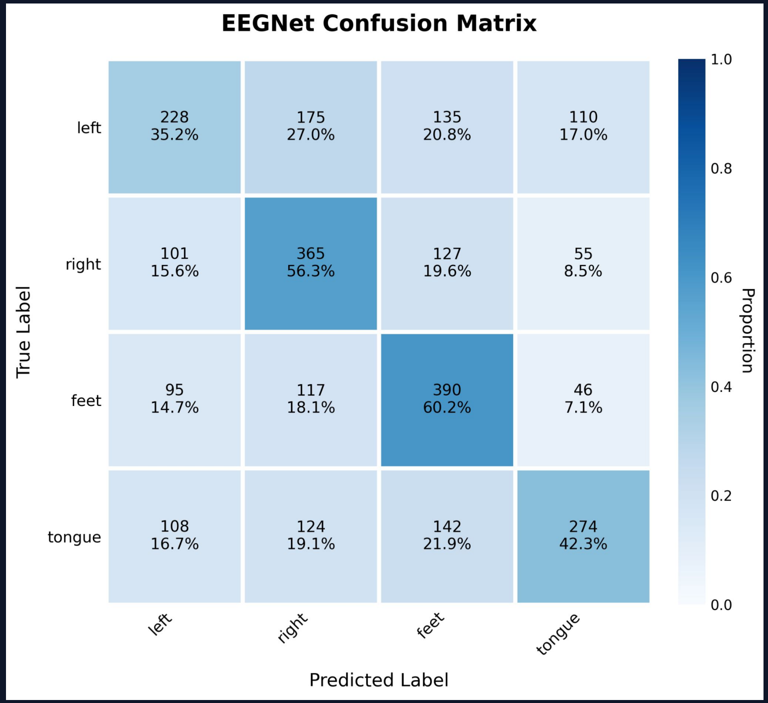
Model	Architecture Type	Trainable Params	Depth
FBCSP + SVM	Signal Proc. + Linear	~1,000 (Non-Deep)	N/A
EEGNet	Compact CNN	~2,500	Shallow (3 blocks)
ATCNet	TCN + Attention	~150,000+	Deep (Multiple TCN layers)

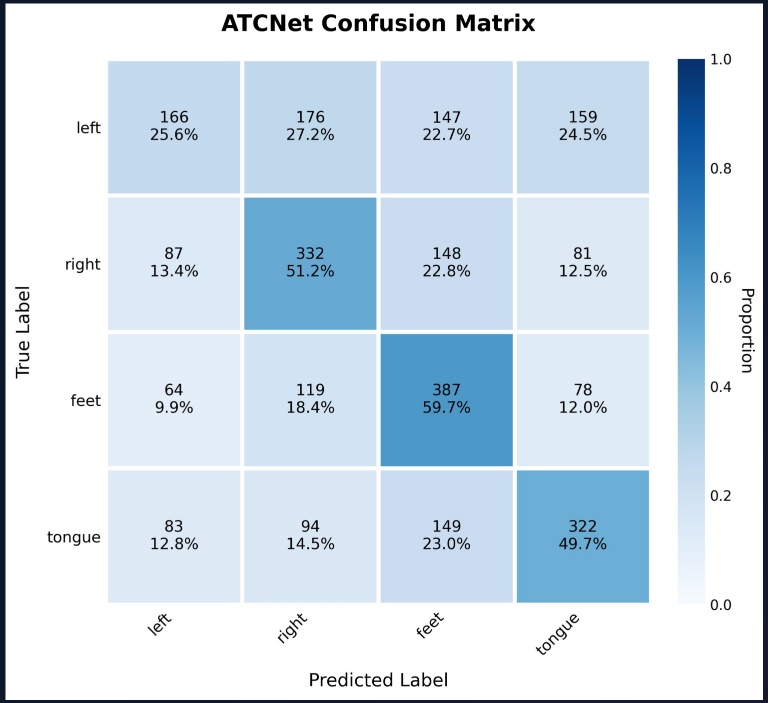
Hypothesis Check:

$$\text{Performance} \propto \text{Complexity ?}$$

EVALUATION RESULTS (LOSO)







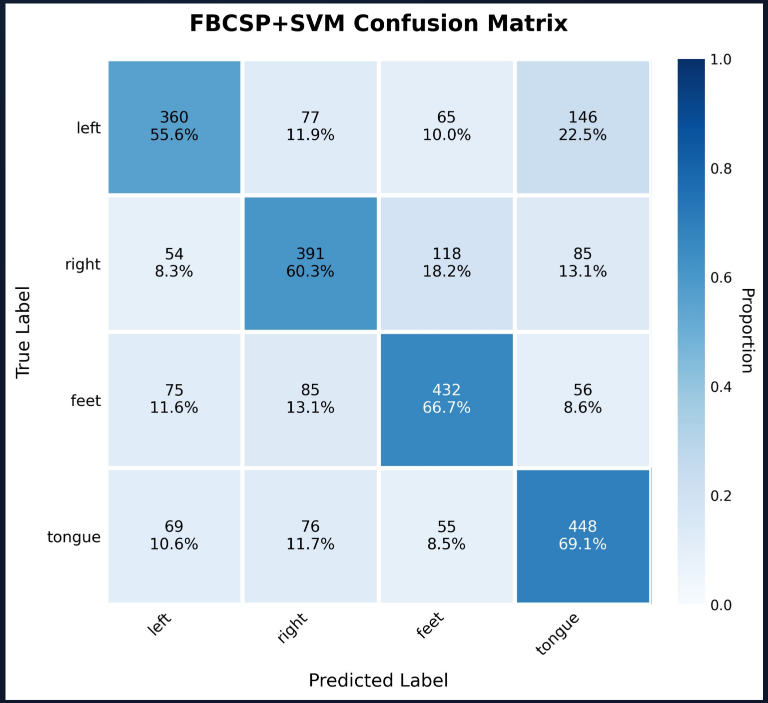
FBCSP + SVM: 38.93% (± 11.65)

EEGNet: 48.50% (± 11.51)

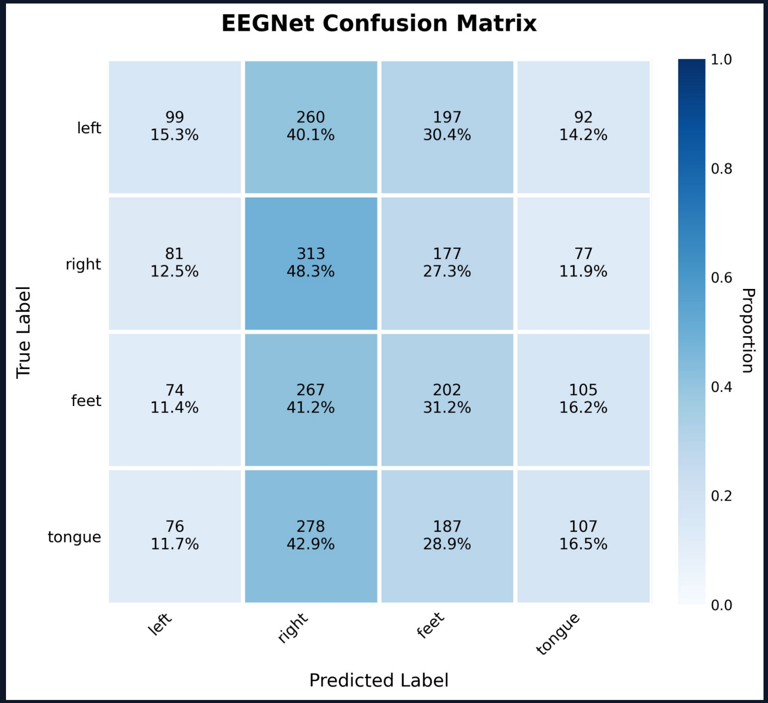
ATCNet: 46.56% (± 11.61)

⚠️ **Note:** ATCNet did ***not*** outperform the lightweight EEGNet in the cross-subject setting.

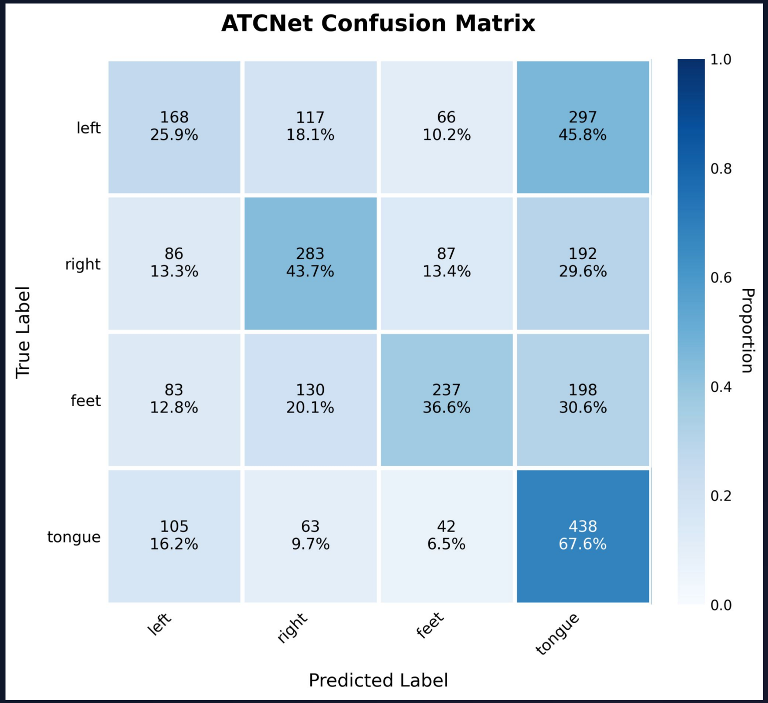
EVALUATION RESULTS (WITHIN SUBJECT)



FBCSP + SVM: 62.92% (± 13.27)



EEGNet: 27.82% (± 2.81)

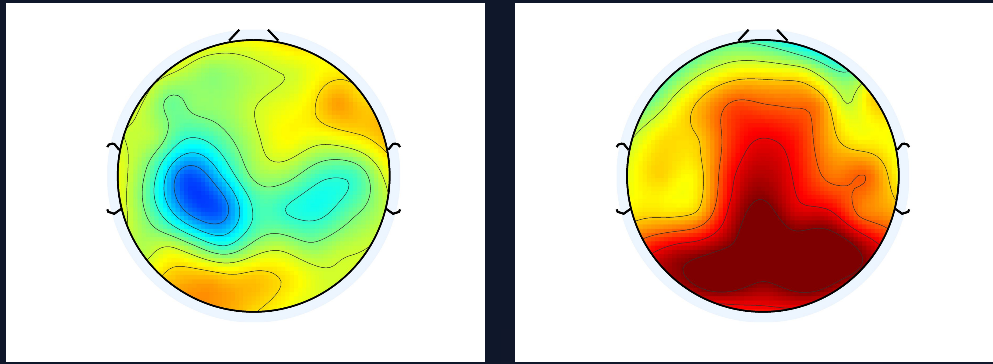


ATCNet: 43.44% (± 15.73)

⚠️ **Note:** FBCSP + SVM far outperformed the two deep learning models.

OBSERVATIONS - GENERALIZATION

- **Variance:** For generalization, it is important to be able to pick up the general features that are associated with a variety of individuals. Deeper networks are usually better at doing this, as they are excel in nonlinear classification problems. On the other hand, the FBCSP+SVM may have struggled to find the components to maximize the variance across subjects, whose MI features can be very different.
- **Lack of Data:** With only 9 subjects, the deep models lack the diversity needed to leverage their parameter count, which may explain why a) ATCNet does not outperform EEGNet and b) the accuracies of the deep learning models do not outperform the FBCSP+SVM model by a greater margin.

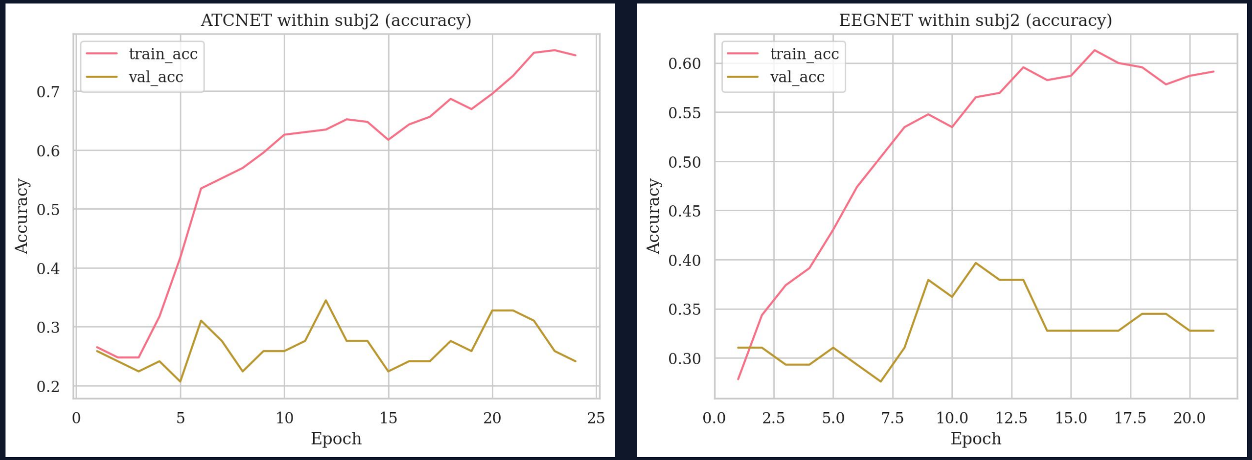


Topographies are during "Motor Imagery" across different subjects, demonstrating the high variance across subjects.

OBSERVATIONS - PERSONALIZATION

- **Overfitting:** The higher complexity/depth of the deep learning models likely led it to memorize specific noise rather than distinguishing the actual subtle patterns in the presence of limited data.

- **Subtlety:** While there is a high variance across subjects, motor imagery within subjects can often be very similar with subtle differences. Based on what we learned about how PCA works, the use of CSP can break down these components and possibly better classify them compared to deeper models, which may explain its performance.



The deeper networks both showed signs of overfitting, as training accuracies continue to increase even while validation accuracies stagnate.

CONCLUSION

GENERALIZATION VS PERSONALIZATION

Deeper/complex networks are better at learning generalized features from large and diverse datasets, with more data improving this learning all the more.

When subtle differences exist within classes in the *same* subject, simple linear models such as SVMs are fully capable of differentiating these, particularly when data is limited.

BOTH? BOTH, BOTH SOUND GOOD.

Though the best case scenario for BCI would be to have a giant corpus of EEG data for each individual, recording EEG data is tedious and very time/effort consuming. Training deep models on large datasets of varying MI EEG data can help with generalization, but may not always outperform a linear model trained on only the particular subject.

An ideal system could combine these together, using deep learning to find general features, which inform the classification of a model that can pick up subtle differences.

REFERENCES

[1] Lawhern, V. J., et al. (2018). EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of Neural Engineering*.

[2] Altaheri, H., et al. (2022). Physics-informed attention temporal convolutional network for EEG-based motor imagery classification. *IEEE Transactions on Industrial Informatics*.

[3] Brunner, C., et al. (2008). BCI Competition IV Dataset 2a. *Graz University of Technology*.

[4] Ang, K. K., et al. (2008). Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface. *IJCNN*.