



AIMS, NIST 2024

Integrated Data Science and Computational Materials Science in Complex Materials

Dilpuneet (DP) Aidhy

Department of Materials Science and Engineering
Clemson University
July 18, 2024

Contributors: Hossein Mirzaee¹, Ramin Soltanmohammadi¹, Nathan Linton², Jacob Fischer², Gaurav Arora³, Serveh Kamrava¹, Pejman Tahmesabi¹ and Dilpuneet Aidhy²

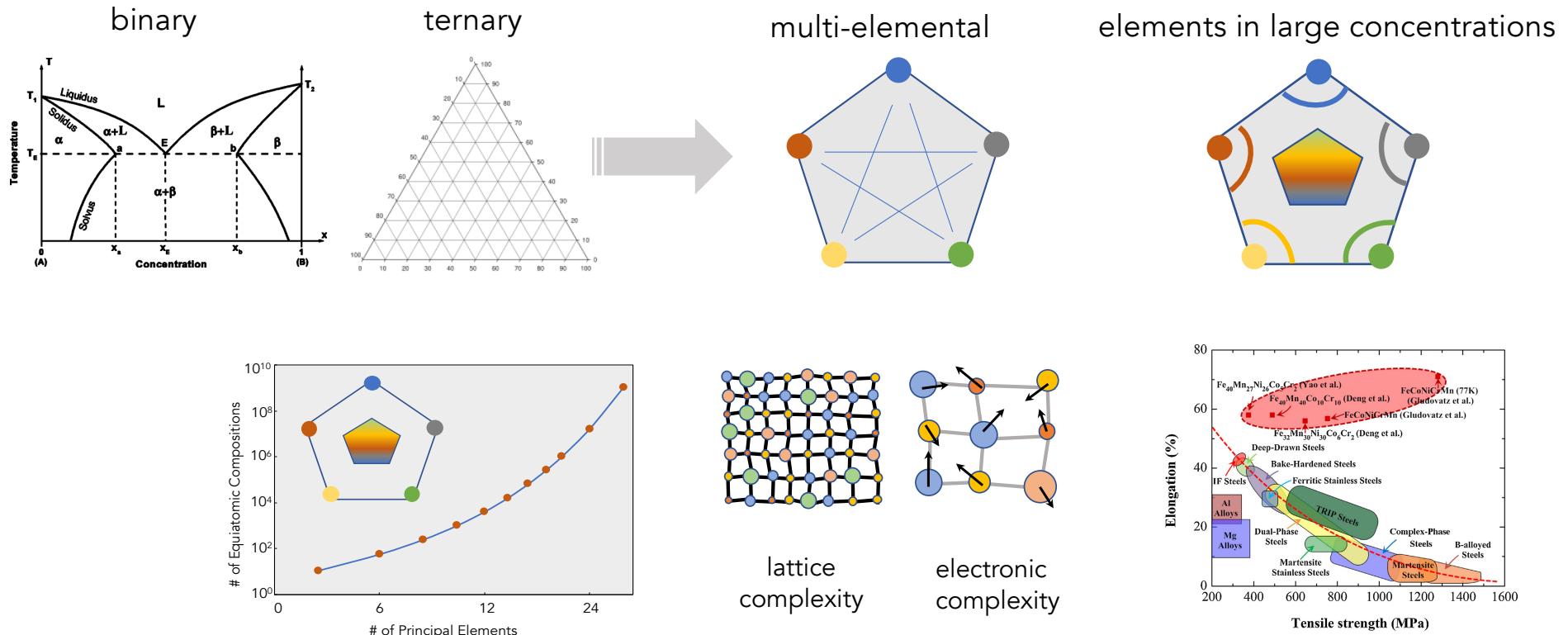
¹Department of Petroleum Engineering, Colorado School of Mines, Golden CO, 80401

²Department of Materials Science and Engineering, Clemson University, Clemson, SC 29634

³Department of Mechanical Engineering, University of Wyoming, Laramie, WY (now at Fermi Lab)



Mindset change: Multi-elemental Materials²



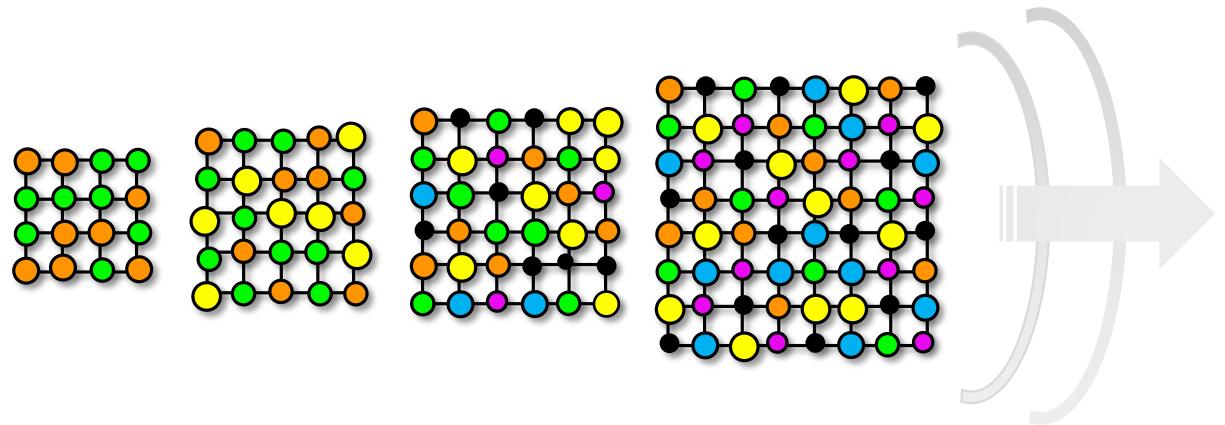
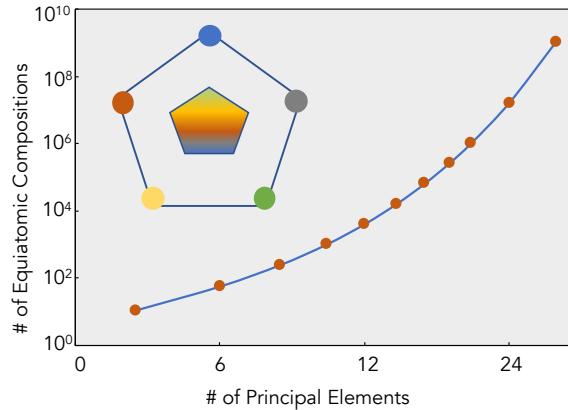
Multi-elemental materials – breaking the Hume-Rothery limits

Field initiated by Yeh and Cantor separately in 2004

Oh et al., Nat Comm. 10, 2090 (2019)

Large Phase Space: Enormous Challenges

3



Challenges:

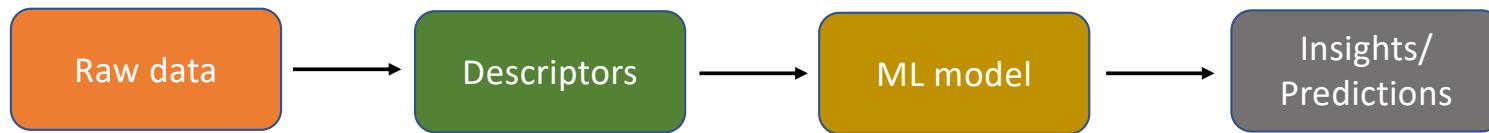
- How to explore phase stability?
- How to explore compositions with interesting properties?
- How to overcome the large computational expense?

From computational viewpoint, conventional methods are limited in handling this challenge

Integrated data-science methods help overcome these challenges

Machine Learning and The Conundrum of Descriptors

- Descriptors contain physics
- They are critical to ML model development
- Ascertaining correct descriptor(s) is a vexing problem
- **Textbook descriptor values change in an alloy**



Common descriptors

Valence electron configuration
Electronegativity
Melting temperature
Atomic radius
Ionization energy
Bulk modulus
Nearest neighbors
Bond type and angles
...

The Fundamental Descriptor: Charge Density

The Kohn-Sham DFT is built on charge density

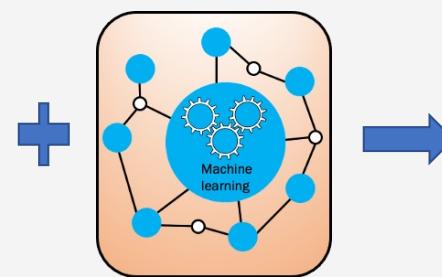
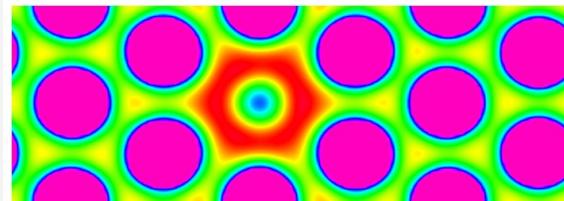
The charge density describes the system completely. All properties emerge from it

Hohenberg-Kohn theorem: Total energy of a many-body electron system is a functional of charge density

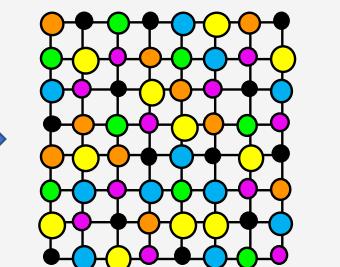
$$E = f[n(r)]$$

$$E = \int dr n(r)V_n(r) - \sum_i dr \varphi_i^*(r) \frac{\Delta^2}{2} \varphi_i^*(r) + \frac{1}{2} \iint dr dr' \frac{n(r)n(r')}{|r-r'|} + E_{XC}[n]$$

Conceptual idea: Use charge density and feature recognition to predict materials properties



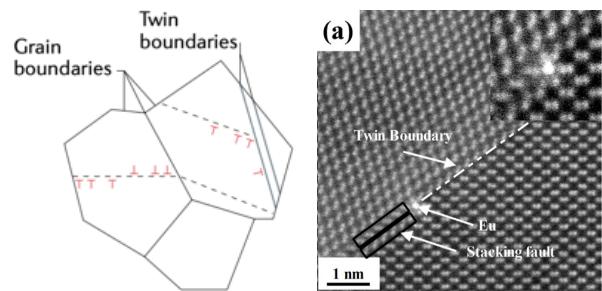
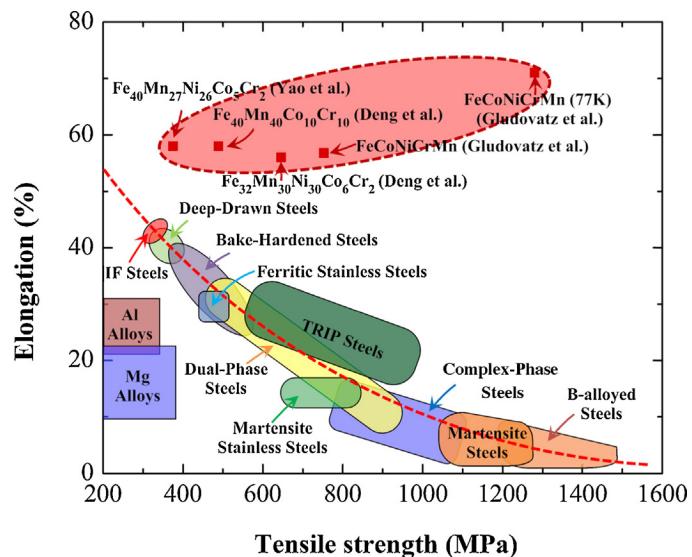
Charge density as the input data



predict properties

Breaking the strength vs ductility tradeoff

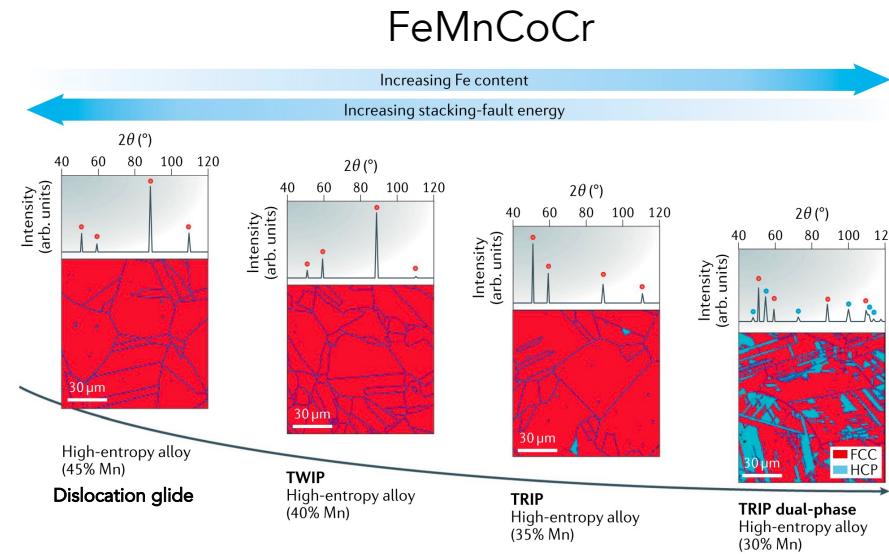
6



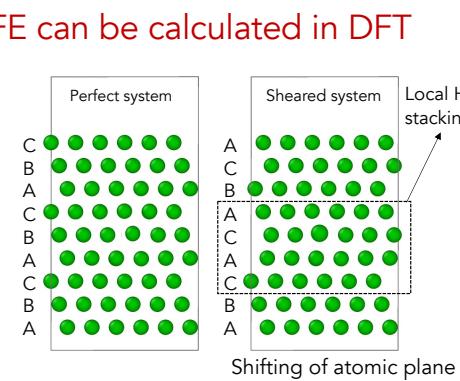
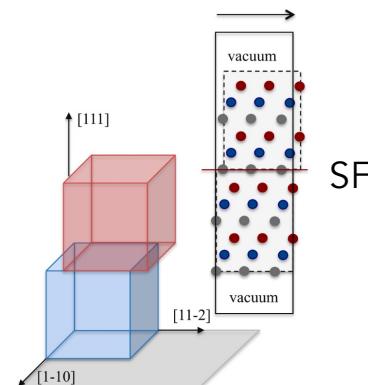
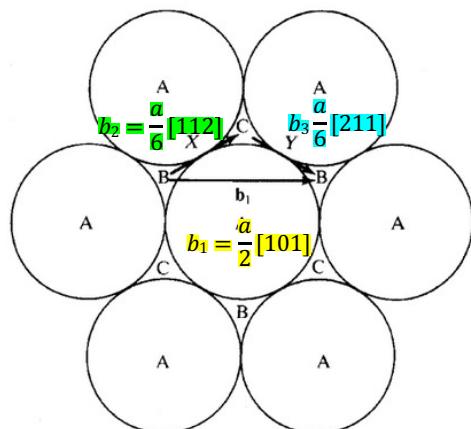
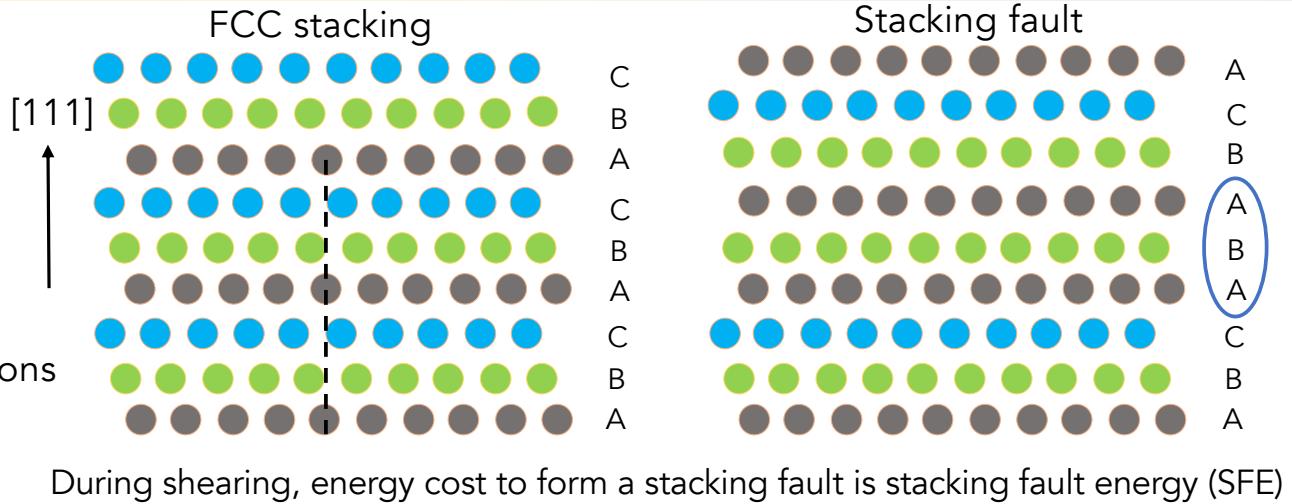
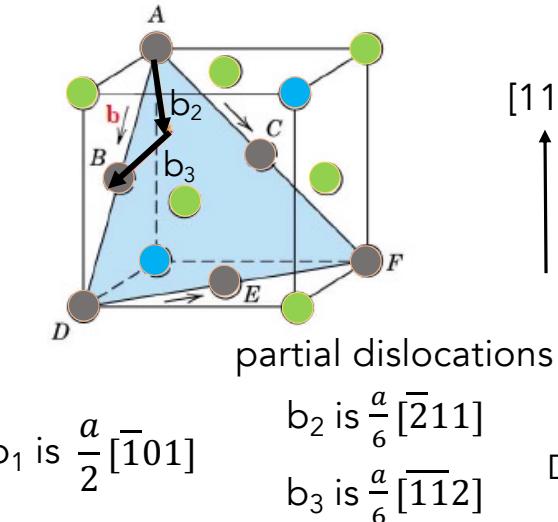
- Low SFE enables twin formation
- Composition and SFE are correlated
- But composition-SFE mechanistic understanding is absent
- SFE tailoring via composition is not yet possible

Ye, Y., Wang, Q., Lu, J., Liu, C., & Yang, Y. (2016). Materials Today, 19(6), 349-362.
George, E., Raabe, D., & Ritchie, R. (2019). Nature Reviews Materials, 4(8), 515-534.

Li, Z. et. al., Mater. Sci. Eng. A, 648 (2015)

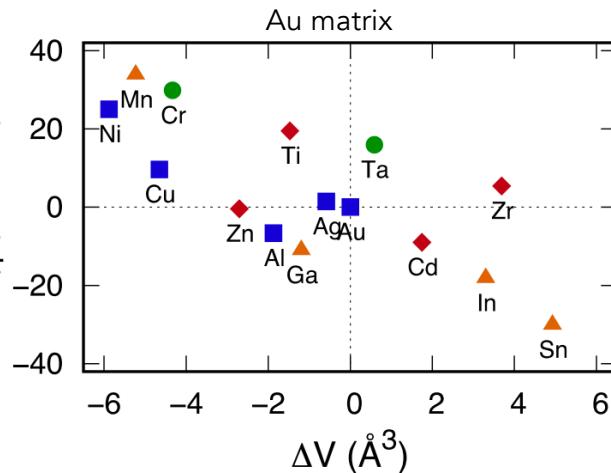
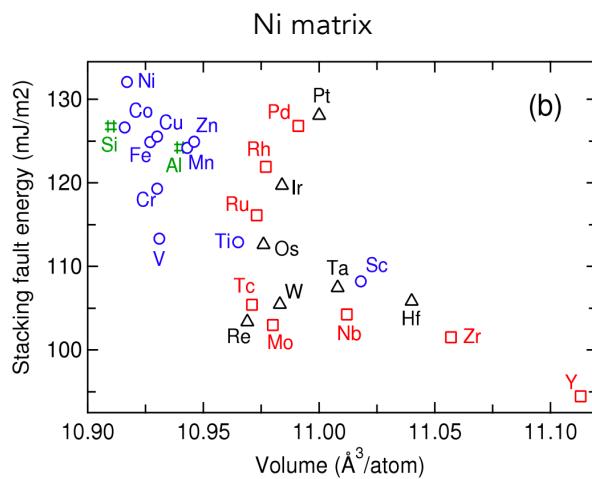


Stacking fault in FCC structure



Literature: 'Descriptors' of SFE

Atomic volume



Charge density



Non-spherical distribution hinders the shear deformation, resulting in larger SFE

- In general, larger atomic volume lowers SFE
- Shape of charge density affects the SFE magnitude

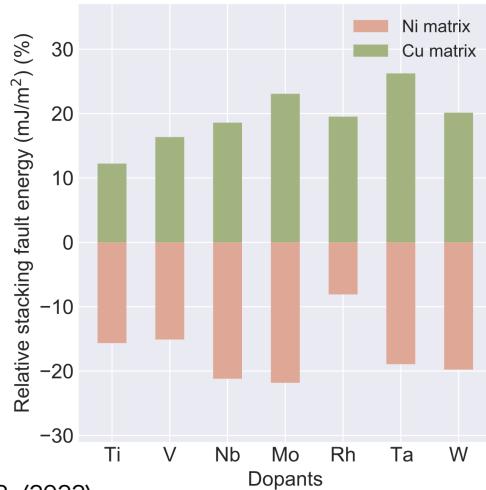
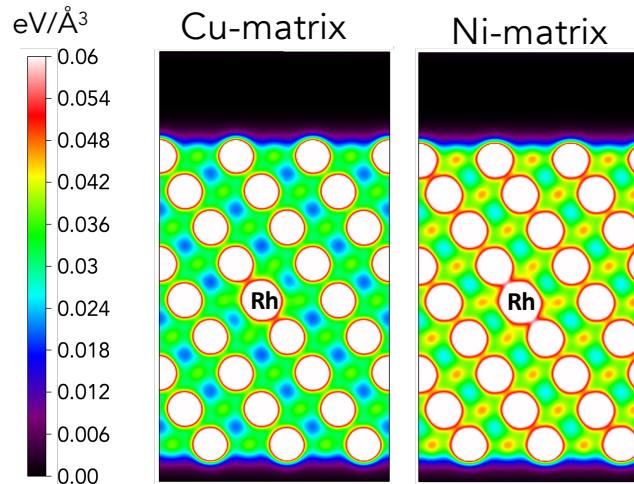
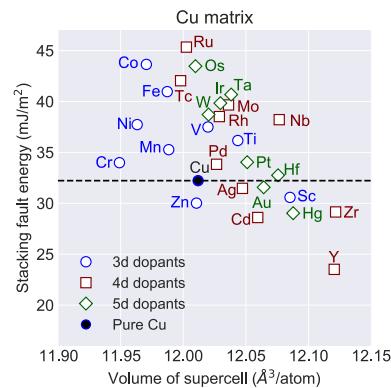
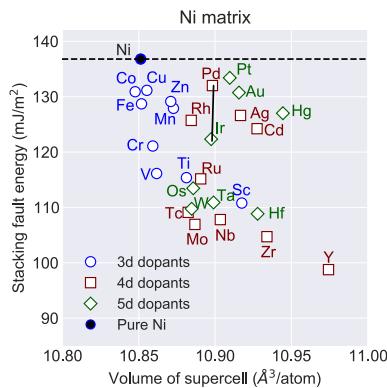
Zhang, Q, et. al., *Intermetallics* 29, 2012

Ogata, S. et. al. *Science* 298 (2002)

Goyal et. al. CMS, 188, (2021)

Shang, S. L. et al. *J. Phys. Condens. Matter* 24, 1–14 (2012).

Effect of atomic volume and charge density on SFE⁹



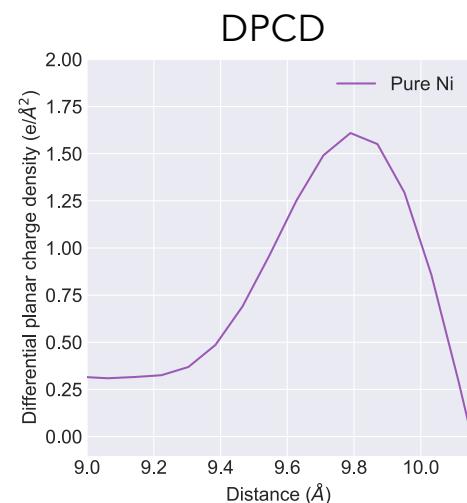
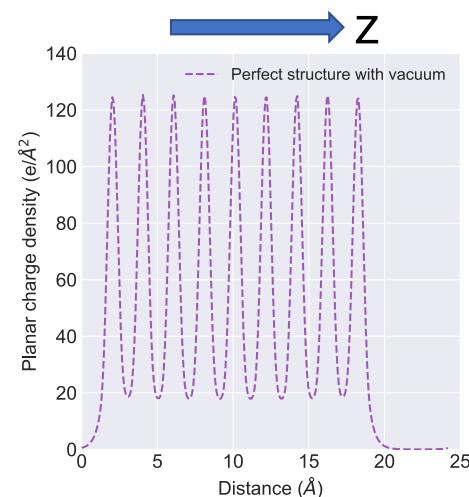
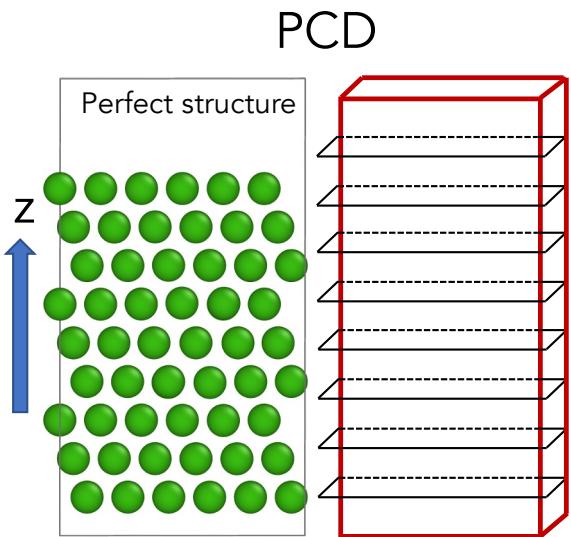
Arora.. Aidhy, JAP, 132, 22, (2022)

- Contrasting effect of the same dopant in Ni and Cu
- Visually, charge density cannot be used as standalone variable for predicting SFE

Charge density before and after shearing

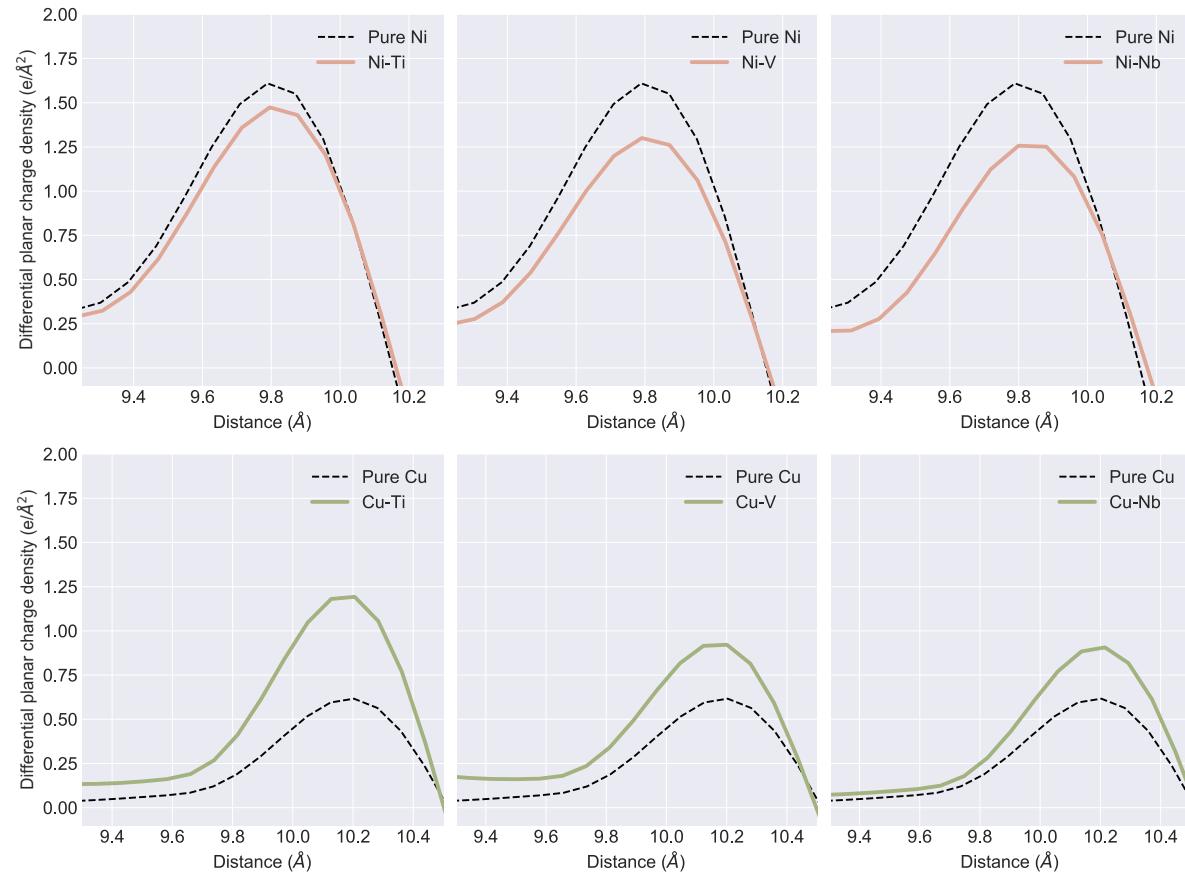
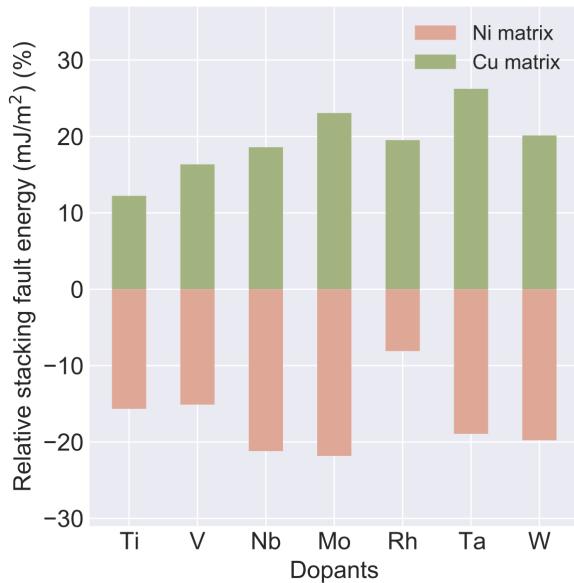
PCD: Planar charge density

DPCD: Differential planar charge density

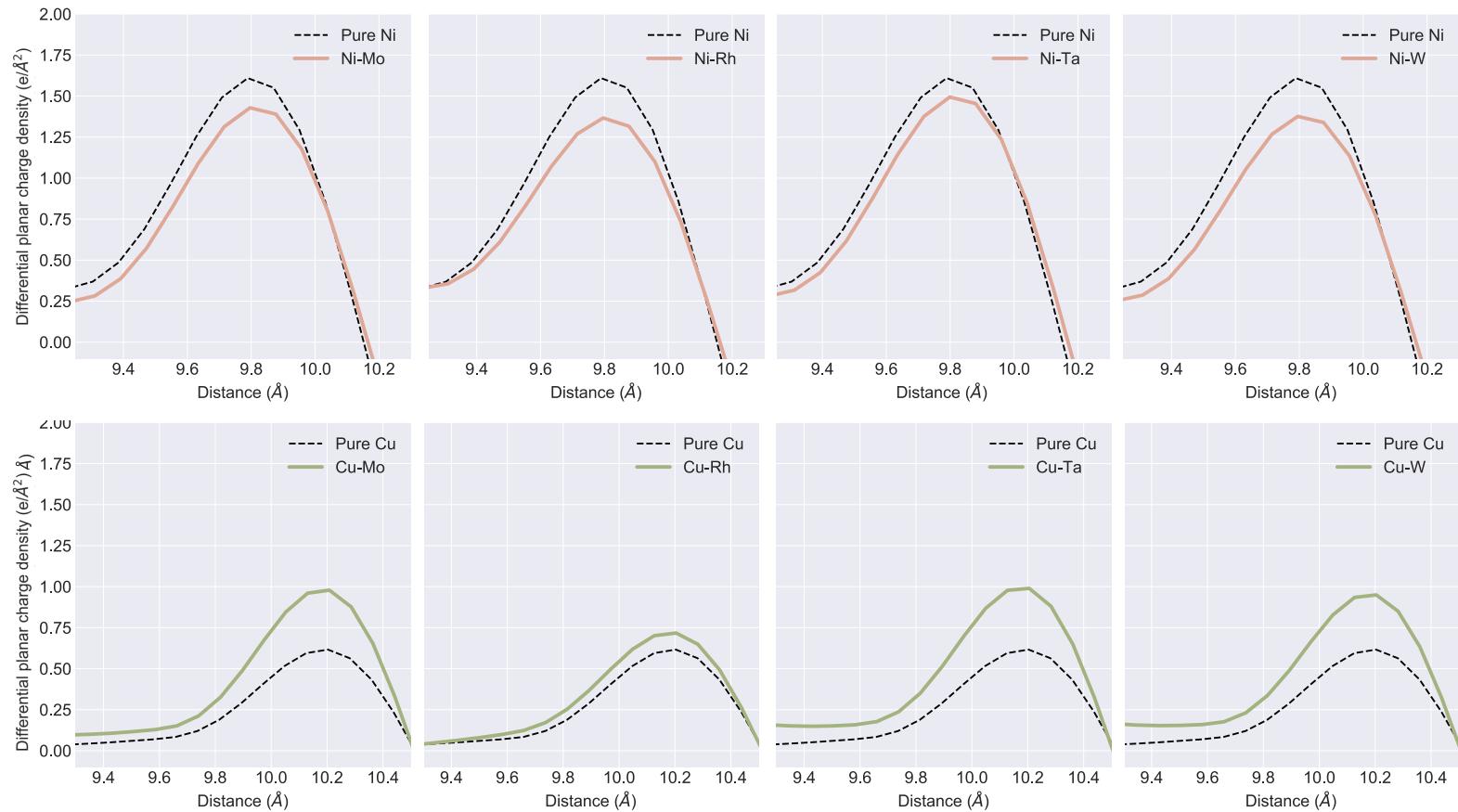


- Peak in DPCD curve indicates distortion in the charge density
- Higher distortion implies higher change in SFE

Results: DPCD comparison in Ni and Cu matrices¹¹



Results: DPCD comparison in Ni and Cu matrices



Higher DPCD results in higher SFE and vice-versa

Arora.. Aidhy, JAP, 132, 22, (2022)

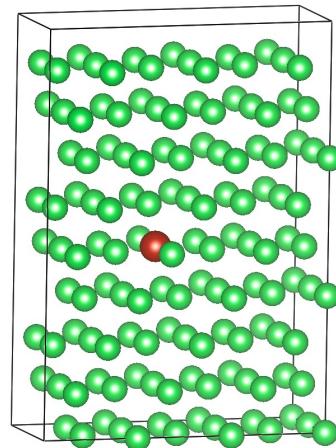
ML prediction on binary alloys

Descriptors: PCD, DPCD, VEC, Volume and Bader charge

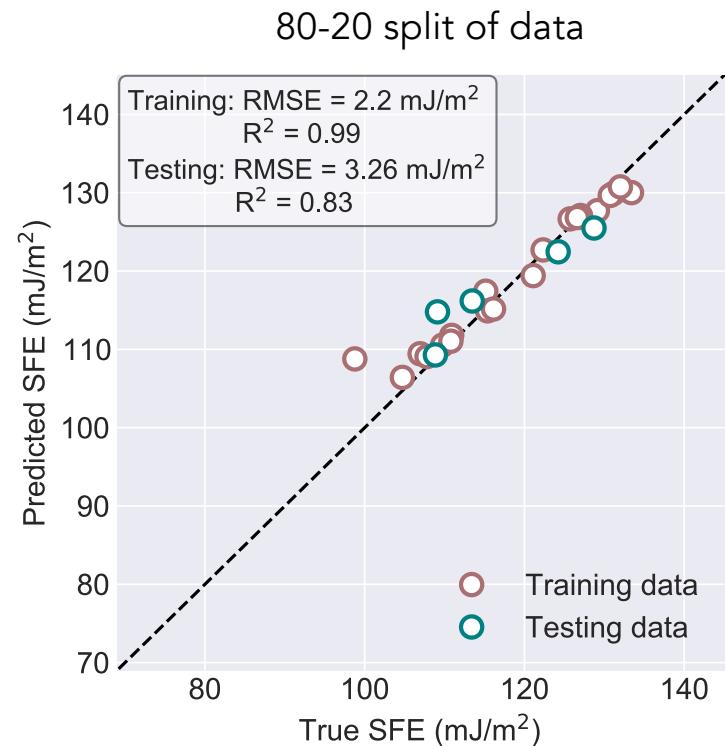
Algorithms tested: Linear regression, **Support vector regressor**,
Random Forest, Gradient Boosting regressor, etc.

Validation tool: RMSE and R²

Single doped binary system

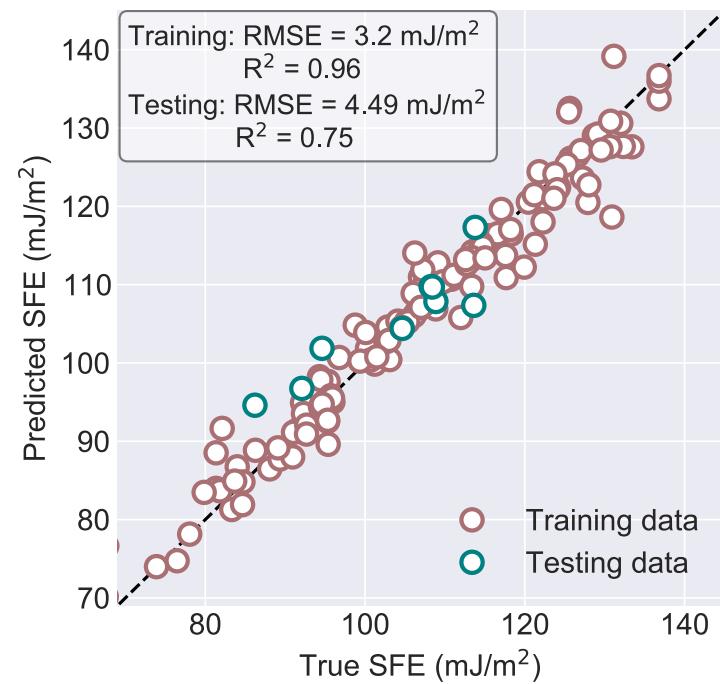
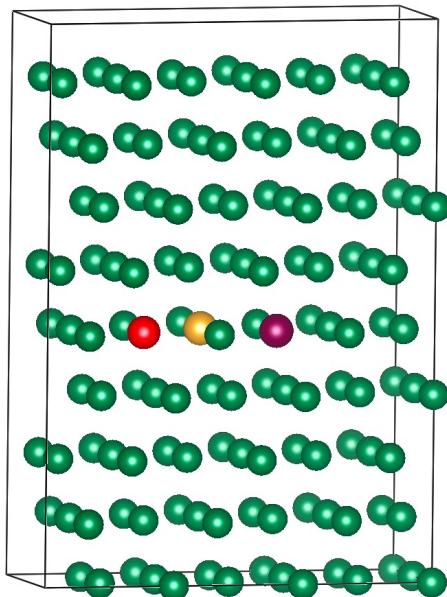


ML is used to predict SFE in dilute binary alloys



ML prediction on quaternary alloys

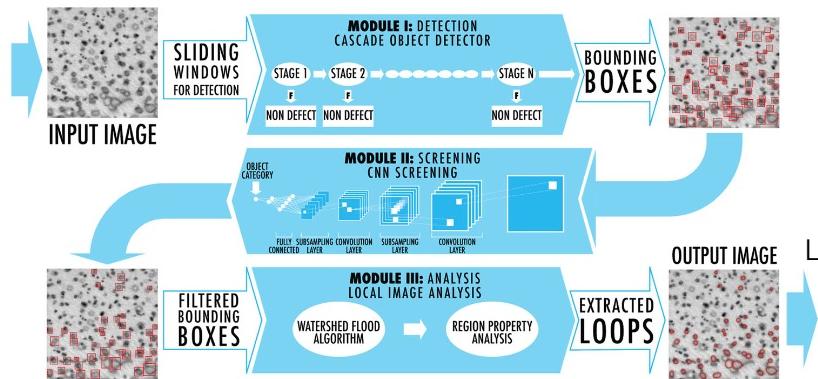
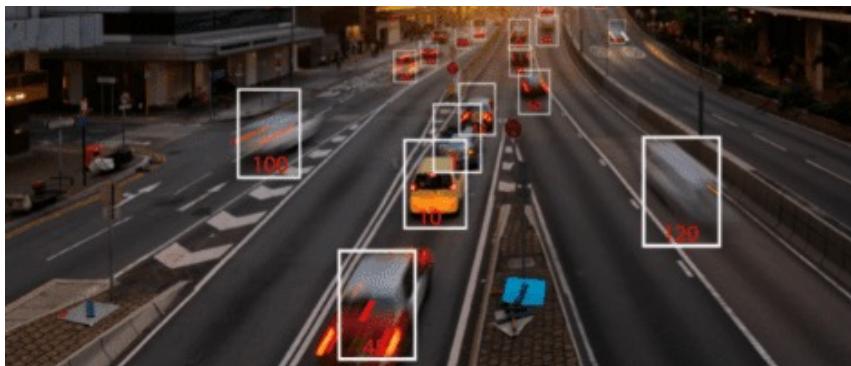
Quaternary



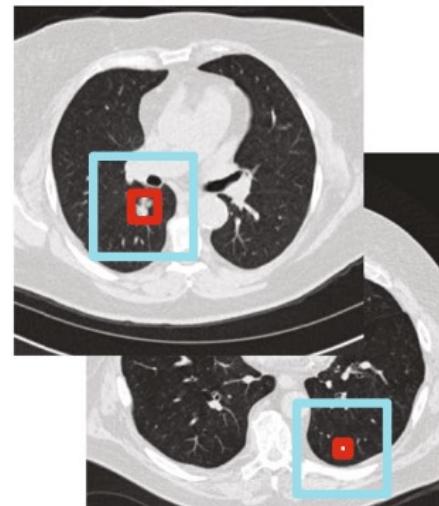
Composition	True (mJ/m ²)	Predicted (mJ/m ²)
NiFeCrCo (1)	104.65	106.60
NiFeCrCo (2)	108.87	109.10
NiFeCrCo (3)	108.27	109.04
NiCrWV	86.21	87.23
NiScCrCo	92.08	98.58
NiScCrCu	94.61	101.18
NiCuCrCo	113.78	118.20
NiCuMnCr	108.43	109.59
NiMnCrCo	113.61	107.08

ML predicts SFE for quaternary alloys by training on binaries and ternaries

Computer vision and Engineering



- Malignancy probability
- LUMAS risk bucket
- Cancer localization



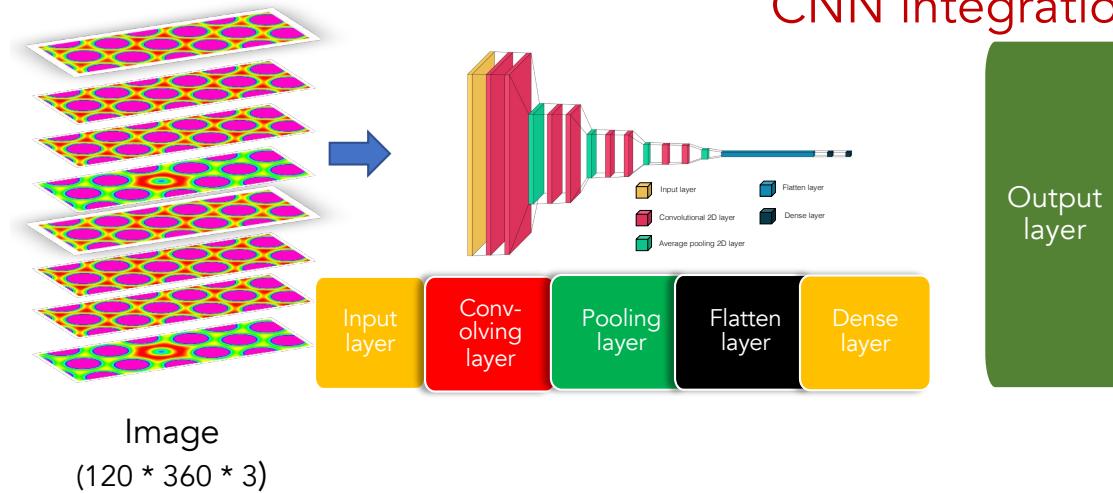
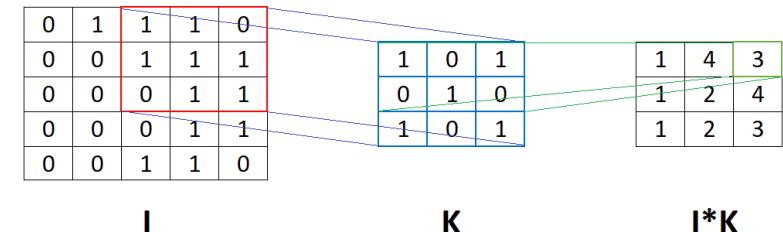
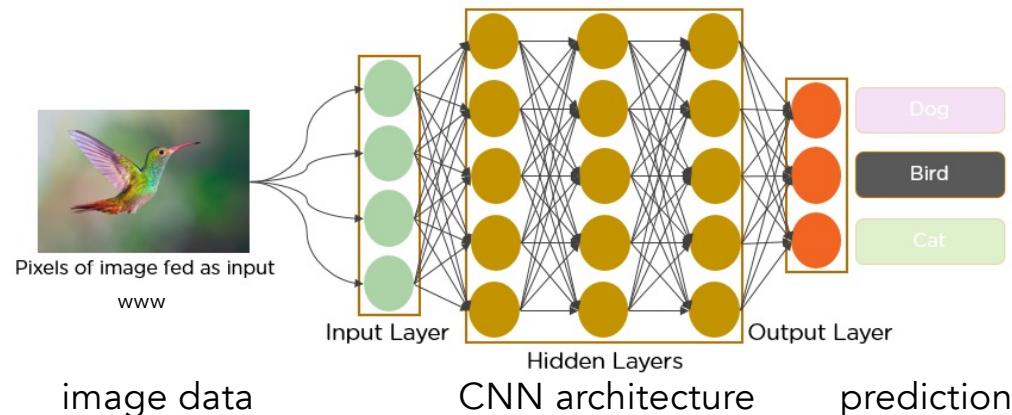
<https://doi.org/10.1038/s41591-019-0447-x>

Li et al., npj Comp Materials, 4,36 (2018)

Learn the patterns from images to predict future possibilities

Feature Recognition using Convolutional Neural Network (CNN)

16



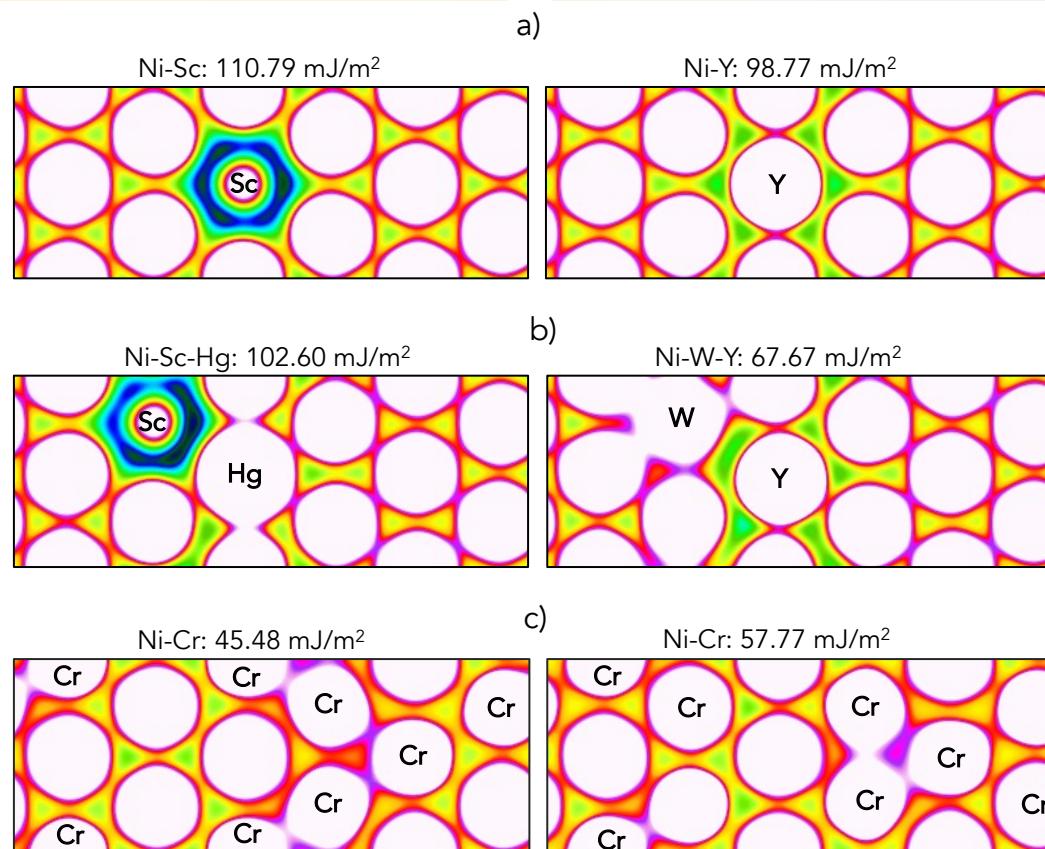
```
#Making CNN model
def build_and_compile_model():
    input_layer = keras.Input(shape = (shape_of_images_data_all[1], shape_of_images_data_all[2], shape_of_images_data_all[3]))
    x = layers.Conv2D(filters = 32, kernel_size = (2,2), padding = 'valid', activation = 'relu')(input_layer)
    x = layers.Conv2D(filters = 64, kernel_size = (2,2), padding = 'valid', activation = 'relu')(x)
    x = layers.AveragePooling2D(pool_size = (2,2), padding = 'valid')(x)
    x = layers.Conv2D(filters = 96, kernel_size = (2,2), padding = 'valid', activation = 'relu')(x)
    x = layers.Conv2D(filters = 128, kernel_size = (2,2), padding = 'valid', activation = 'relu')(x)
    x = layers.AveragePooling2D(pool_size = (2,2), padding = 'valid')(x)
    x = layers.Conv2D(filters = 160, kernel_size = (2,2), padding = 'valid', activation = 'relu')(x)
    x = layers.Conv2D(filters = 192, kernel_size = (2,2), padding = 'valid', activation = 'relu')(x)
    x = layers.AveragePooling2D(pool_size = (2,2), padding = 'valid')(x)
    x = layers.Conv2D(filters = 256, kernel_size = (2,2), padding = 'valid', activation = 'relu')(x)
    x = layers.Flatten()(x)
    x = layers.Dense(16, activation = 'relu')(x)
    outputs = layers.Dense(1)(x)
    model = keras.Model(inputs = input_layer, outputs = outputs)

    model.compile(
        loss = keras.losses.mse,
        optimizer = keras.optimizers.Adam(learning_rate = 0.01),
        metrics = ['MeanSquaredError'])

    return model
```

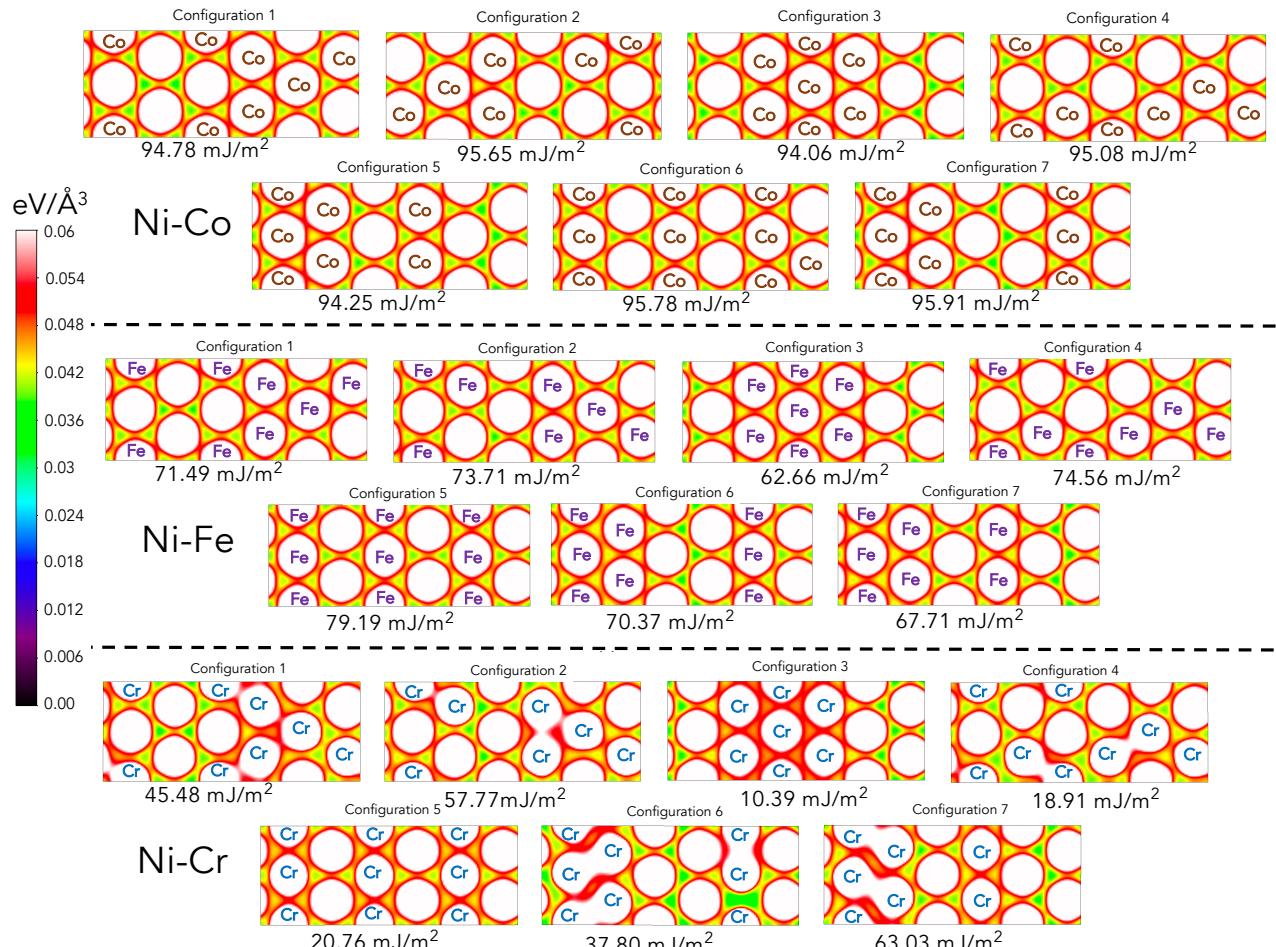
Arora...Aidhy, Materialia, 26 (2022)

Charge Density and SFE Variations

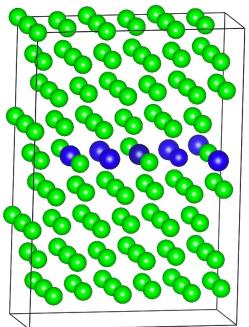


Charge density distribution and SFE are correlated

Variation in SFE and charge density distortion – binary alloys



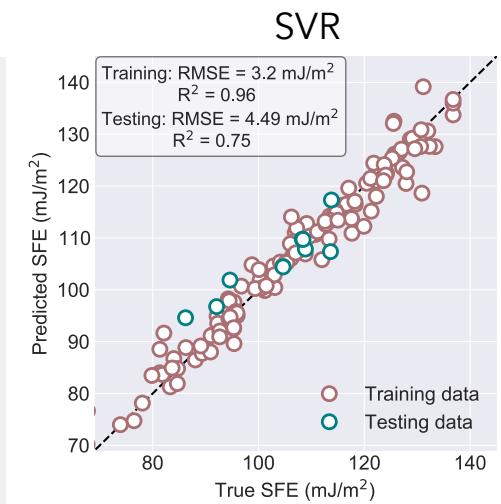
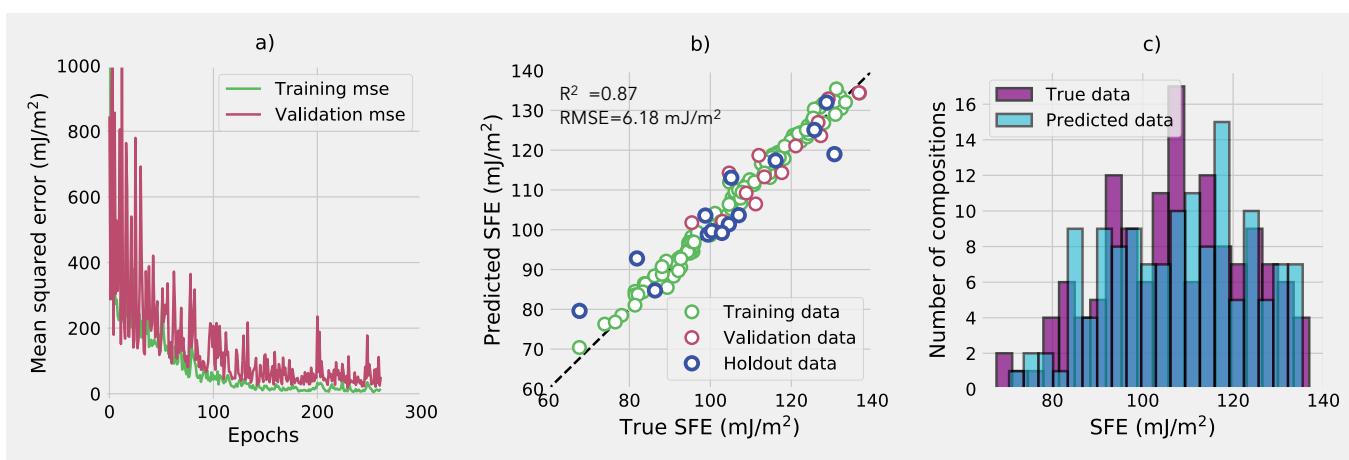
Alloy system	SFE Variation (mJ/m^2)
Ni-Co	1.85
Ni-Fe	16.93
Ni-Cr	52.64



- Diverse NN environments may still yield similar SFE (e.g. NiCo)
- Greater charge distortion results in larger SFE variation (e.g. NiCr)
- SFE variation possibly emerges from charge density distortion, and not necessarily from different NN environments

SFE prediction from CNN

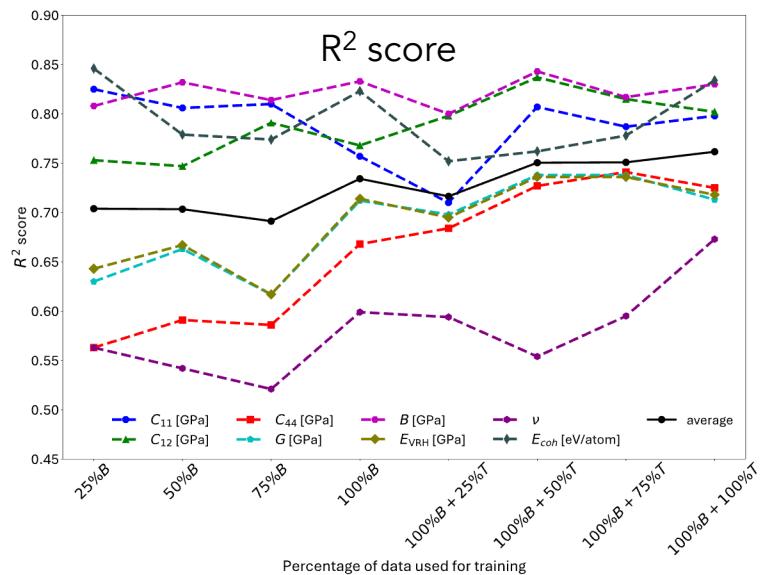
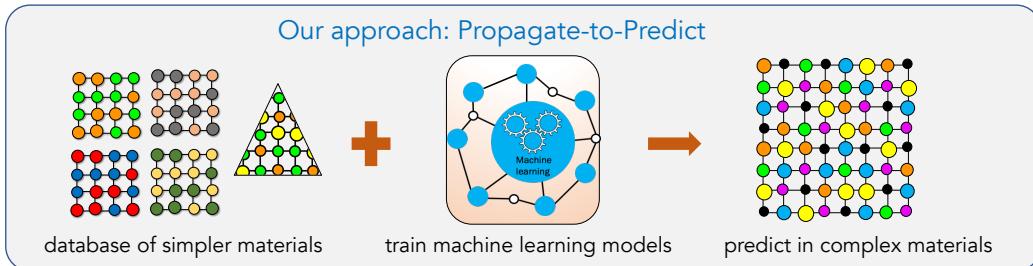
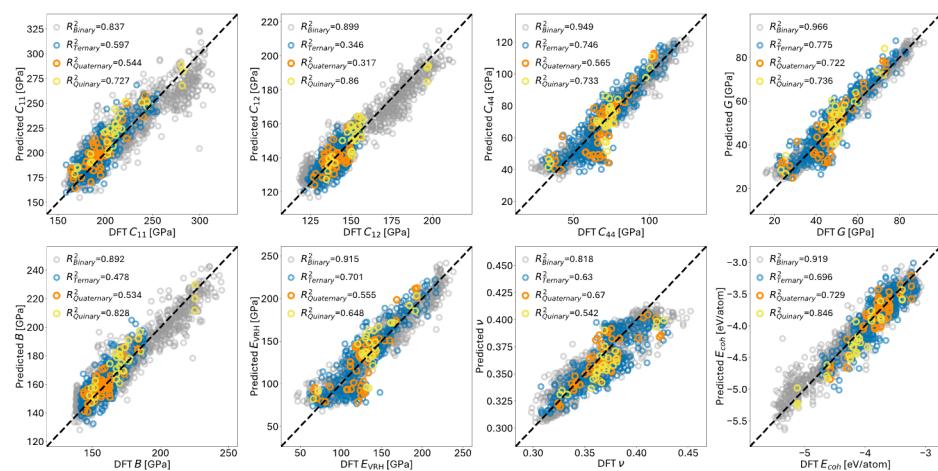
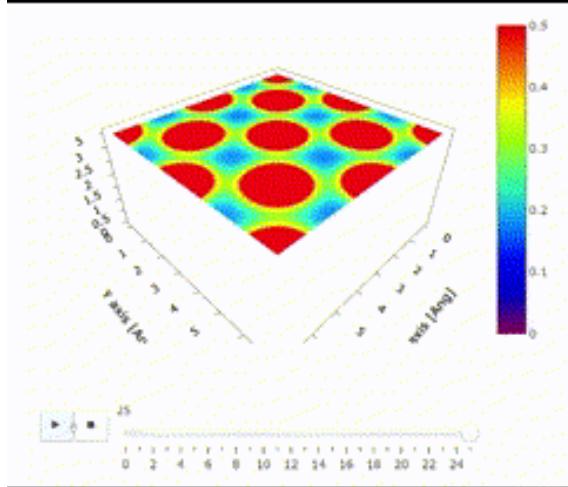
Model is trained using 139 data points, including binaries, ternaries, and quaternaries with a split of 80-20



SFE is predicted from charge density distribution using image recognition methods



Elastic constants from Charge Density





Materials
Computation &
Data Science

MC•DC

Concluding Remarks

- Only charge density is used as a descriptor for SFE predictions in ML model
- Feature recognition CNN is applied to predict properties
- Charge density approach is agnostic to type of property
- Some limitations: consistent pixel quality and system size

Acknowledgements



CDS&E: “Charge-density based ML framework for efficient exploration and property predictions in the large phase space of concentrated materials”
under award number 2302763.



Materials
Computation &
Data Science

MC•DC

