Rollout policy   SL policy network   RL policy network   Value network

p_π p_σ p_ρ v_θ

Neural network

Policy gradient

Classification   Classification

Self Play

Regression

Data

Human expert positions   Self-play positions
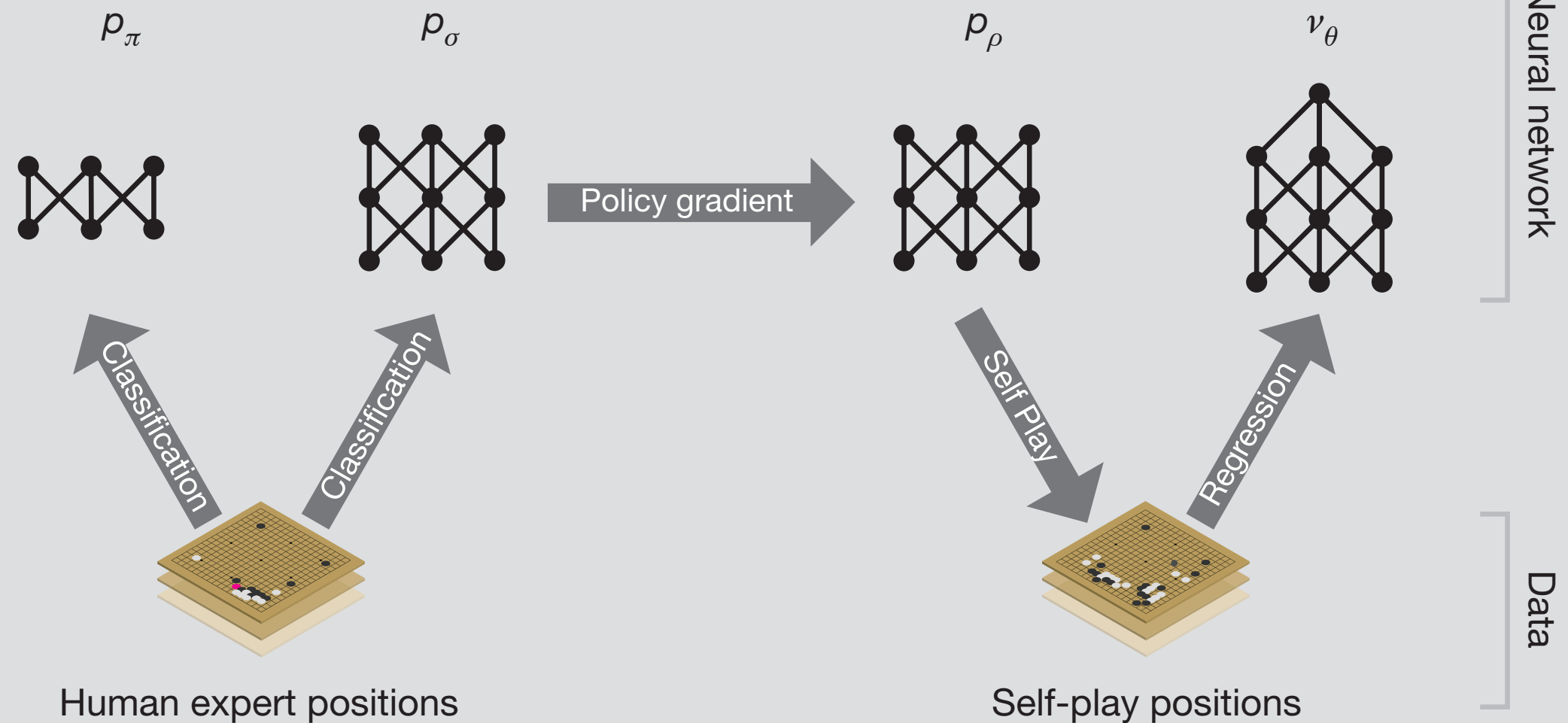
- SL Policy network

- 13 layer CNN

- Accuracy 57% VS 44.4%

- Fast Rollout Policy

- local feature & linear model

- Accuracy 24.2% & 2us VS 3ms

- RL Policy network

- Self-play 30 million times

- Value network

- self play data set

- add

Extended Data Table 4 | Input features for rollout and tree policy

Feature	# of patterns	Description
Response	1	Whether move matches one or more response pattern features
Save atari	1	Move saves stone(s) from capture
Neighbour	8	Move is 8-connected to previous move
Nakade	8192	Move matches a *nakade* pattern at captured stone
Response pattern	32207	Move matches 12-point diamond pattern near previous move
Non-response pattern	69338	Move matches 3×3 pattern around move
Self-atari	1	Move allows stones to be captured
Last move distance	34	Manhattan distance to previous two moves
Non-response pattern	32207	Move matches 12-point diamond pattern centred around move

Features used by the rollout policy (first set) and tree policy (first and second set). Patterns are based on stone colour (black/white/empty) and liberties (1, 2, ≥ 3) at each intersection of the pattern.