It has been shown that the brain uses feature detection, in its visual pattern recognition task.

Many researches have attempted to capture the brain's pattern recognizing capability in abstract models.

In their attempts, some have tried to remain faithful to the biological principles underlying the functioning and organization of the brain.

Others borrowed from the brain the most important principles and tried to cast them in any model that could be demonstrated to work.

The model in this paper follows the latter approach.

The work done by Fukushima will first be examined, representing earlier models of its class.

A new model, which attempts to overcome their limitations, will then be described.

Finally the paper will conclude with a description of future improvements to the model.

Models which loosely follow the brain's architecture, have done so based on the following elementary principles:

a) The neuron (as a threshold element) is the building block of those models.

b) The neuron's output can be viewed as a boolean corresponding to on and off activation states, or as a positive (bounded) real number representing the activation rate of the neuron.

c) The pattern recognizing network is layered.

d) Each layer contains feature detecting cells with increasing level of conceptual complexity, the higher the layer level.

Many models in the neural network literature applied those principles.

Focus will be centered on Fukushima's model, because it is the most elaborate and has been shown to work with a (relatively) large retina of 128x128 pixels.

Fukushima's Neocognitron model has layers with two levels each.

The lower level is made up of groups of template matchers.

The higher level neurons take their inputs from groups at the lower level that recognize the same object at different positions.

The net effect is feature detectors tolerant of displacement and slight distortion.

Fukushima's neocognitron suffers the following problems.

First, hardware is not amiable to large scale implementations: in one case, a simple 19x19 retina, 4 layer network implementation required over 40 thousand cells, excluding non-responding ones.

Second, each learned template is duplicated in many positions after being trained in only one position: this is problematic for hardware implementations as well as being biologically implausible.

This model is based on the premise that instead of taking the feature detector to the pattern (multiplying the number of feature detectors), we bring the pattern to the feature detector (multiplying the pathways.)

As the example of the neocognitron shows, duplicating complex feature detectors is costly, in terms of number of cells.

What we hope to achieve is a reduction in the overall number of cells.

The retina is thus divided into several marginally overlapping receptive fields (RF's) of uniform size.

Simple hard-wired operators provide a many-to-one mapping from the RF area to the feature detector.

Those operators are divided into classes.

For instance, on the lowest layer, the classes are displacement, scaling, and rotation.

On higher layers, the classes include positional and set operators.

Each class has its variations within each RF, depending on where the operator maps from, and the degree of the mapping.

For example, the displacement class has variations which corresponds to the direction and the amount of the displacement.

On the next level, within a layer, feature detectors take their input from the output of the operators weighed by the optimal feature pattern.

The output of those feature detectors represents the degree of success with which a particular operator maps into the optimal feature.

From this large pool of feature detector outputs within a receptive field, the best variation and degree for each class is selected.

Then, the optimum values across the layer from each RF are combined to choose the best class of operators.

This choice represents the consensus as to which class of operators best maps into some feature.

The consensus is then fed back to lower levels, allowing each RF to reset its own image of the retina according to the new transformation.

Notice that while the application of the operator class is enforced upon the layer, each RF implements the transformation in a way that generates optimal mapping.

After one class is selected within a layer, the class is then inhibited allowing another class to win in a second round.

The process is then repeated until a threshold of desirable outcome is exceeded.

The feature with best degree of success, can thus be said to have been recognized.

Having recognized a feature for each RF on the first layer, the output of the first layer is fed into the second layer.

A similar process of transformation and recognition is carried out in the second layer.

Finally, on the topmost layer, a feature detector which conceptually represents an object is selected and the whole process terminates.

Simulating this model necessitated additional hardware that was not originally envisioned.

While its design premise is simple, the number of cells required to implement it is proportional to the number of RF's, the number of classes, the number of variations within a class, and the number of features within each layer.

The model does not lend itself to a nice and simple mathematical model to support it, and mathematical properties of the model are not practical

to implement.

For example, while it is mathematically sound to say that two features are different if one can not be generated from the other by applying any sequence of operators in any order.

This property taken to the extreme is not practical to use in order to incorporate self-organization into the model.

While that the model can be augmented with learning rules that change the weights on the feature detectors, it fails to address how a whole new class of operators can be learned.

The model has been shown in practice to fail under certain circumstances: the wrong sequence of transformations is applied leading to either faulty recognition or no recognition at all.

This is a result of the lack of communication between neighboring RF's.

A feature detection-based model was presented that was designed to address the limitations of previous models.

The model was developed from an innovative idea.

Although the model achieved its objective of lesser overall hardware complexity, it had few limitations of its own.

Currently, work is in progress on a new model.

This model incorporates communication between neighboring RF, coupled with hill climbing techniques to pick the best transformation to apply within an RF image.

In effect, this will result in a reduction in the number of pathways as only few transformations will be implemented at a time.